



**UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL**

**DISEÑO E IMPLEMENTACION DE UN SISTEMA DE INFORMACION DE APOYO A
LA GESTION Y PROYECCION DE DEMANDA EN UNA AEROLINEA**

MEMORIA PARA OPTAR AL TITULO DE INGENIERO CIVIL INDUSTRIAL

FELIPE ANDRES MARAMBIO ABARCA

PROFESOR GUÍA:
NICOLAS JADUE MAJLUF

MIEMBROS DE LA COMISIÓN:
HERNAN CARDENAS HERMOSILLA
ANTONIO HOLGADO SAN MARTIN

SANTIAGO DE CHILE
JUNIO 2016

RESUMEN DE LA MEMORIA PARA OPTAR
AL TÍTULO DE: INGENIERO CIVIL INDUSTRIAL
POR: FELIPE MARAMBIO ABARCA
FECHA: 06 DE JUNIO, 2016
PROF. GUÍA: SR. NICOLAS JADUE.

DISEÑO E IMPLEMENTACION DE UN SISTEMA DE INFORMACION DE APOYO A LA GESTION Y PROYECCION DE DEMANDA EN UNA AEROLINEA

La industria aérea de transporte de pasajeros está inmersa en dinámicas de constante cambio, es muy sensible a las situaciones económicas globales y a las contingencias internacionales. En Latinoamérica, es una industria aún en desarrollo, tendiendo a tener demandas crecientes y oportunidades de nuevos mercados, a diferencia de otras regiones, donde existe un nivel de madurez mayor.

El presente informe da cuenta del trabajo de memoria, el cual tiene por objetivo diseñar e implementar un sistema de información de apoyo a la toma de decisiones táctico-comerciales dentro de una aerolínea, con especial énfasis en las decisiones referidas a la proyección de pasajeros, proceso clave que permite coordinar acciones de corto y mediano plazo en los equipos comerciales y lograr capturar oportunidades.

El trabajo se justifica sobre la base que no existe un sistema de información ni definición homogénea que apoye corporativamente procesos comerciales de corto y mediano plazo, estos procesos se definen en base a muchos reportes y fuentes información, pero que no siempre contienen datos de bases equivalentes e indicadores con una misma definición, aumentando los costos de coordinación. Por lo tanto, es necesario establecer un sistema y proceso claro que permita entregar información de alto nivel para la toma de decisiones.

Para lograr el objetivo planteado es necesario realizar una serie de actividades: Definir los requerimientos de las áreas que utilizarán el sistema, seleccionar las bases de datos, consensuar una definición de los indicadores, realizar la extracción, transformación y carga de estos datos. Finalmente, aplicar una metodología de proyección de demanda para agregarla, posteriormente, a un reporte de información y definir los procesos para su actualización.

Algunas de las dificultades que aparecen, es la gran cantidad de información, y las múltiples fuentes donde puede ser obtenida, lo que complejiza la extracción y análisis de datos. Otra dificultad, es la unificación de los indicadores. Dado que los análisis que se hacen con los datos son múltiples, hay que estudiar cuales son las métricas adecuadas y como construirlas.

Los resultados son prometedores, en la implementación piloto, se logró disminuir el tiempo de los analistas de negocio dedicado a obtener datos en un 11%, las proyecciones obtenidas en mercados de mayor ingreso tienen un margen de error cercano al 10%, que se traduce en mejores decisiones en el corto y mediano plazo.

A Isidora, por permitirme soñar
A mi familia.

Agradecimientos

Al profesor Nicolás Jadue por sus consejos y disponibilidad en todo momento.

A mi familia por acompañarme en mi periodo educativo y entregarme todas las herramientas para crecer.

Agradezco a LAN por abrirme los espacios de trabajo.

Agradezco especialmente a Isidora Vielma por su apoyo incondicional e iluminar mi mundo.

TABLA DE CONTENIDOS

1. INTRODUCCIÓN	1
1.1 Industria Aérea y LATAM.....	1
1.2. Descripción y justificación del proyecto	3
1.3. Objetivos.....	4
1.3.1. Objetivo General	4
1.3.2. Objetivos Específicos	4
1.4. Alcances	5
1.5. Resultados esperados	6
2. PROBLEMA DE NEGOCIO.....	8
2.1 Revenue Management.....	8
2.2 Problemática de definiciones	11
2.3 Sistemas de información de LATAM.....	11
3. METODOLOGÍA.....	13
3.1 Entendimiento de la problemática y requerimientos principales del sistema	13
3.1.1 Dinámica del negocio general.....	13
3.1.2 Requerimientos principales	14
3.2 Diseño lógico de métricas y atributos	14
3.2.1. Levantamiento de métricas y atributos generales	14
3.2.2. Métricas específicas de proyección de demanda.....	14
3.3 Diseño del modelo de fuentes de información y procesos.....	15
3.3.1 Identificación fuentes de información	15
3.3.2 Diseño de procesos.....	15
3.3.3 Elección de herramientas.....	16
3.4 Visualización	16
3.4.1 Diseño de visualización	16
3.4.2 Elección de plataforma visual.....	17
3.4.3. Medición de resultados, métricas específicas de proyección.....	17
4. MARCO CONCEPTUAL.....	18

4.1	Conceptos comerciales y métricas del negocio de transporte	18
4.1.1	Gestión comercial y sistemas de distribución	18
4.1.2	Conceptos geográficos y de vuelo	19
4.1.3	Conceptos relacionados a la venta y canales	21
4.1.4	Indicadores comerciales y de rentabilidad	23
4.2	Modelos y métodos cuantitativos para el pronóstico de demanda.....	25
4.2.1	Modelos predictivos.....	25
4.2.2	Modelo básico económico de oferta y demanda en una aerolínea	26
4.2.3	Series de Tiempo	27
4.3	Modelos de procesos y bases de datos.....	35
4.3.1	Modelamiento de procesos de negocio.....	35
4.3.2	Modelos de bases de datos	36
4.3.3	Procesos ETL.....	37
5.	DESARROLLO DEL TRABAJO	39
5.1	Selección de principales métricas y atributos del negocio	39
5.1.1	Análisis de métricas y atributos	41
5.2	Proceso ETL	46
5.2.1	Extracción de datos	46
5.2.2	Transformación de datos y carga	47
5.3	Requerimientos de negocio asociados a la proyección	48
5.3.1	Información y procedimiento	48
5.3.2	Performance herramienta visualización	49
5.3.3	Proceso de proyecciones	49
5.3.4	Proyección de demanda y reservas históricas	50
5.3.5	Pickups Multiplicativos en las reservas	53
5.3.6	Matrices de anticipación.....	54
5.3.7	Modelos de proyecciones.....	55
5.4	Solución como sistema de información.....	56
5.4.1	Elección de herramienta visualización y funcionamiento	56
5.4.2	Presentación visual de la herramienta	58

5.4.3 Arquitectura final del sistema	65
6. ANALISIS DE LOS RESULTADOS	67
7. CONCLUSIONES Y COMENTARIOS FINALES	69
7.1 Mejoras y trabajos futuros.....	71
8. BIBLIOGRAFIA	73
9. ANEXOS	74
A. Software proyección demanda	74
B. Pentaho data integration	84
C. PALO.....	86

INDICE DE FIGURAS

Figura 1: Flota LATAM	2
Figura 2: Distribución por país empleados LATAM	2
Figura 3: Modelo de funcionamiento macro esperado	7
Figura 4: Proceso de comercialización [11].....	9
Figura 5: Modelo de trabajo actual táctico RM [11]	10
Figura 6: Decisiones gestión comercial.....	18
Figura 7: Ejemplo definición segmentos y tramos	20
Figura 8: Ejemplo de series de tiempo	27
Figura 9: Matriz de indicadores principales	41
Figura 10: Modelo métrica pasajero reserva	43
Figura 11: Modelo métrica pasajero industria.....	43
Figura 12: Relación ticket y cupón	44
Figura 13: Modelo métrica pasajero venta, ingreso y tarifa media	45
Figura 14: Capa integración	47
Figura 15: Canal de reserva	49
Figura 16: Cuadro comparativo semanas.....	51
Figura 17: Grafico reservas mercado SCL-BUE.....	52
Figura 18: Grafico reservas mercado SCL-BUE por POS.....	52
Figura 19: Cuadro de Pick-Ups	53
Figura 20: Cuadro por anticipación y pickups.....	54
Figura 21: Proceso cálculo proyección.....	55
Figura 22: Matriz resultado pickups.....	56

Figura 23: Modelo de visualización	57
Figura 24: Herramienta / Vista POS	58
Figura 25: Herramienta / Vista agrupación OD POS	60
Figura 26: Herramienta / Vista ODPOS.....	62
Figura 27: Herramienta / Vista Sub ruta	64
Figura 28: Vista gráficos.....	65
Figura 29: Modelo final.....	66

1. INTRODUCCIÓN

1.1 Industria Aérea y LATAM

La industria aérea es un mercado altamente competitivo, debido principalmente a los bajos costos de entrada al negocio. Esto ha permitido la proliferación de distintas líneas aéreas durante los últimos años, lo que ha provocado una caída en las tarifas y una ampliación del espectro socioeconómico de los clientes a los cuales se les presta el servicio. Por otro lado, Latinoamérica presenta un incremento anual en la demanda mayor que en otras regiones, dado el nivel de madurez del mercado. Sin embargo, actualmente se ha visto opacado por la situación económica, principalmente de Brasil.

LATAM Airlines Group, es un grupo de aerolíneas actualmente conocidas como LAN y TAM, que se creó en junio de 2012, tras la fusión de estas últimas, llegando en ese momento a ser el grupo de aerolíneas de mayor valor bursátil del mundo. Hoy el escenario es diferente, llegando a un valor de MMUS\$ 3 180 (enero 2016) ¹, explicado por el contexto de la región, en la cual sus competidores también presentan caídas considerables.

LATAM Airlines Group se divide en dos líneas de negocio: carga y pasajeros, este último, es en el que se enfoca el presente trabajo y se divide a su vez en tres negocios: doméstico (vuelos dentro Chile, Perú, Argentina, Ecuador, Colombia, Brasil y Paraguay); regional (vuelos dentro de Sudamérica) y largo alcance (Norteamérica, el Caribe, Oceanía y Europa).

Desde el año 2000 pertenece a la alianza OneWorld (elegida regularmente como la mejor alianza de líneas aéreas²), a la que pertenecen entre otras, aerolíneas como; American Airlines, British Airways e Iberia. LAN y TAM desempeñan un rol protagónico dentro de la alianza ya que es el principal actor dentro de la región. Gracias a esta alianza LATAM Airlines Group puede ofrecer más de 300 destinos en el mundo.

LATAM posee una flota de aviones relevantes, la principal de la región. A marzo 2016 la flota de aviones de pasajeros se compone de:

¹ Yahoo Finance. <http://finance.yahoo.com>. Enero 2016

² Skytrax Airlines world ranking. http://www.worldairlineawards.com/Awards/worlds_best_airline_alliance.html. 2015 Awards

Flota de aviones pasajeros	
Avión	Total
Airbus A319-100	50
Airbus A320-200	154
Airbus A321-200	36
Airbus A330-200	10
Airbus A350-900	1
Boeing 767-300	38
Boeing 777-300 ER	10
Boeing 787-800	10
Boeing 787-900	7
Total	316

Figura 1: Flota LATAM

En cuanto a los detalles de la organización, pertenecen a la estructura alrededor de 50,413 empleados³. Estos se distribuyen en diversos países:



Figura 2: Distribución por país empleados LATAM

³ Memoria resultados anual 2015.

Además, LATAM Airlines cuenta con un programa de pasajeros frecuente llamado *Fidelidade* para pasajeros TAM y *Lanpass* para pasajeros LAN. En la actualidad estos 2 programas de beneficios se han unido (01 de marzo de 2016) para dar origen a “LATAM Pass” que es el nuevo programa frecuente que combina los participantes de los anteriores programas.

El mayor porcentaje de las ganancias de la compañía proviene del negocio de pasajeros, éste ascendió a MMUS\$ 513 en 2014, correspondiente a ingreso operacional (20,3% más bajo que el 2013), con un margen operacional de 4,1% (0,7% más bajo que 2013).

Es de interés notar que los resultados anuales de la compañía reflejan una gestión estructural de años previos, dado que el producto vendido es un “servicio a futuro” por lo tanto se requiere tener una gestión intensiva al mediano y largo plazo.

1.2. Descripción y justificación del proyecto

Los sistemas informáticos de la compañía almacenan datos de reserva y venta/facturación en distintas fuentes (dependiendo del tipo de venta). Estos datos se registran desde más de 15 años y son usados por diversas áreas de la empresa, sin embargo, no existen capas de indicadores ni definiciones únicas, lo que provoca que las acciones de las diferentes áreas no estén alineadas o estén basadas en datos erróneos que pueden llevar a decisiones equivocadas y descoordinaciones. Además, la tarea de consolidación de la información dada la complejidad de la industria y el tamaño de la empresa provoca que más del 10% del tiempo invertido de los analistas sea dedicado a esta labor.⁴

Frente a esta situación se ha concebido idear un sistema de información que sirva a las diferentes áreas del negocio, que contenga indicadores comunes calculados de una misma forma y además proponga una proyección de demanda de pasajeros, esto como ejercicio para un mejor entendimiento del mercado.

Para contar con indicadores de gestión apropiados es necesario llegar a un consenso sobre cuáles son los indicadores relevantes o significativos, cuál es la forma de cálculo, cuáles fuentes de información son necesarias, cuáles son sus consideraciones, la orquestación de los procesos, la publicación de los indicadores necesarios y bajo qué paradigma tecnológico será aplicado.

Para el caso de la proyección de demanda como apoyo en la gestión se hace relevante una revisión de la literatura actual, entender los modelos, metodologías y analizar bajo qué consideraciones podría ser aceptable la proyección. Se busca consensuar un

número que sirva de apoyo para el entendimiento del mercado y que podría ser modificable por los analistas.

En la literatura actual existen diversos modelos y metodologías que sirven para modelar pronósticos de demanda. Se debe buscar modelos simples que se ajusten a los datos recolectados.

Adicionalmente se cuenta con un sistema *world class* que apoya la gestión táctica comercial del área de *Revenue Management*, llamado PROS. Este sistema en conjunto con otros, soporta las decisiones de tarifas y disponibilidad.

Sin embargo, estos sistemas no responden a las necesidades de información para la gestión, ni tampoco consideran variables adicionales de mercado y la competencia, por lo tanto, los analistas determinan acciones basados en información incompleta o simplemente no utilizan la potencialidad del sistema, gestionando en función de la experiencia lo que puede llevar a conclusiones equivocadas.

El poder analizar con más tiempo la información y basarse en ésta para la toma de decisiones, entender de forma más completa la demanda y tener indicadores consensuados con el resto de las áreas, permite tomar decisiones y ejecutar acciones comerciales para revertir situaciones adversas o aprovechar cualquier oportunidad de negocio que permita generar ingresos incrementales. Por ende, es de real importancia para el negocio contar con este tipo de herramientas ligadas a un proceso establecido y constante que lo sustente, teniendo en cuenta que además estudios demuestran que a medida que más tiempo se utilice para analizar la información, provoca un incremento en los ingresos.

1.3. Objetivos

1.3.1. Objetivo General

- Diseñar e implementar un sistema de información automático de soporte la toma de decisiones tácticas comerciales

1.3.2. Objetivos Específicos

- Consensuar y definir indicadores representativos de la gestión táctica de los negocios de forma coordinada.

- Definir uno o varios modelos de proyección que permitan de forma automática construir una propuesta para los indicadores de demanda de pasajeros.
- Diseñar y unificar los procesos que capturen la información desde las fuentes de datos.
- Construir e implementar una interfaz que muestre esta información, así como también establecer las bases de datos que almacenen la información unificada.
- Medir los resultados y cuantificar los beneficios para el negocio.

1.4. Alcances

El alcance de la memoria es el desarrollo de una herramienta o sistema que soporte la toma de decisiones tácticas de un grupo de actores dentro del negocio. Esto se logra organizando una serie de indicadores de gestión consensuados y definidos entre las áreas y automatizando una proyección de demanda de pasajeros de referencia.

El trabajo de modelamiento se realizará en base al estudio de información histórica de reservas tanto de LAN como también de la industria área. Estos datos provienen de diversas bases de datos que posee la compañía. Posteriormente, se construirá un sistema informático que almacene, actualice y ponga a disposición de los usuarios la información recolectada.

El sistema no pretende en ninguna forma administrar por sí solo las proyecciones de demanda de la compañía, ya que existen factores que no se incluyen en el estudio que son fundamentales al momento de proyectar demanda, tampoco es posible sistematizar la experiencia comercial de los actores del negocio, los lineamientos estratégicos de la compañía, situaciones económicas globales, entre muchos otros factores.

Cabe mencionar que el entregable final de este trabajo no es solo un diseño teórico, sino también una implementación piloto dentro de la empresa, que contenga todos los elementos previamente descritos.

Se considerarán para el estudio de los indicadores y el modelo de datos las bases de información de la compañía, que incluyen tickets y reservas de todos los mercados (denominados típicamente como mercado doméstico, regional y *long haul*).

De esta forma, al abarcar todas las áreas de negocio, el impacto del sistema a implementar es mucho mayor

Para fines de la implementación piloto, la aplicación práctica será exclusivamente para los negocios regionales y *long haul*, en particular para vuelos cuya operación la lleve a cabo LAN Airlines (Código LA).

Los datos para el presente trabajo son de reservas y confirmaciones de pasajes aéreos del año 2008 y 2009. Esto puede ser extensible para años futuros dado que la plataforma es escalable y no tiene limitantes en ese sentido.

El trabajo fue realizado durante el año 2010, pero el contexto industria e información de la compañía han sido revisados para ajustarse al periodo actual.

1.5. Resultados esperados

Se espera tener un sistema de información de gestión de fácil acceso y uso que contenga los indicadores relevantes del negocio unificados, que posea una proyección de demanda a nivel de gestión y que permita apoyar mediante esta información las decisiones táctico-comerciales que recaen en los distintos equipos del negocio, en especial las que tengan que ver con el proceso de proyecciones.

El resultado de los estudios y modelos aplicados permitirán tener información clave y consensuada para el negocio, la cual puede ser utilizada de forma transversal por todos los equipos comerciales. Es decir, se contará con una fuente única de información para la gestión táctica.

El sistema mostrará proyecciones de pasajeros, además de indicadores claves, en un panel actualizado de manera periódica posterior a la recepción de datos de diversas fuentes. Luego, se generarán actualizaciones a la proyección en base a esta información.

El beneficio para el negocio se estima en el ahorro de 1% del tiempo dedicado a la construcción de reportes y obtención de información por parte de los equipos usuarios. Adicionalmente, entre los beneficios cualitativos se encuentran los ahorros en costos de coordinación, lo cual provoca un incremento en el ingreso incremental.

En la figura 3 se observa el modelo macro esperado, donde se tienen los inputs estructurados y no estructurados. Luego de ser procesados por lógicas de negocio unificadas y consensuadas por las áreas usuarias, se tiene una capa de integración que permite disponibilizar periódicamente a los usuarios la información que requieren para el desarrollo de sus actividades diarias.

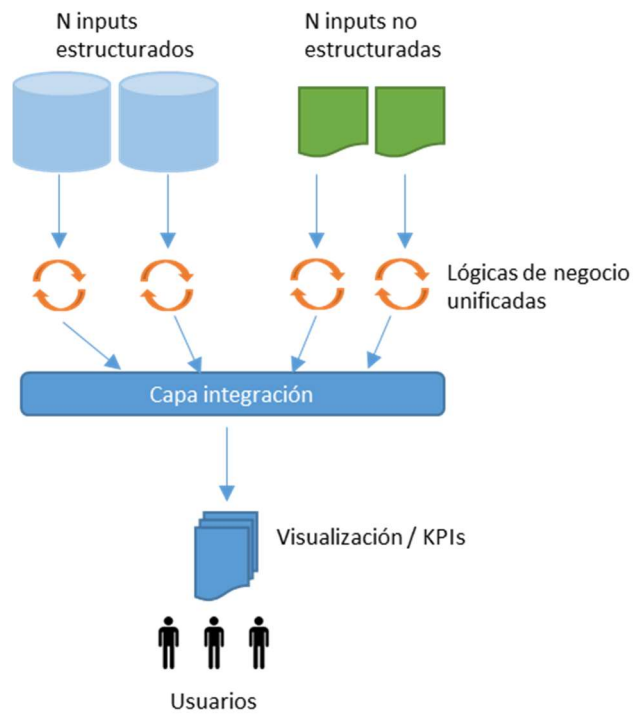


Figura 3: Modelo de funcionamiento macro esperado

2. PROBLEMA DE NEGOCIO

2.1 Revenue Management

Los equipos de *Revenue Management* cumplen un rol relevante dentro del negocio. Esta disciplina es definida comúnmente como el arte y ciencia de predecir la demanda de los clientes a nivel de micromercado y optimizar los precios y disponibilidad de los productos. El principal objetivo que persigue el R.M. es vender el producto correcto al cliente preciso en el momento y precio adecuados para maximizar los ingresos netos de la compañía. Esta maximización de rentabilidad se logra en un ciclo dinámico de procesos de pronósticos de demanda, optimización de precios y disponibilidad de productos, todo esto en función de una demanda esperada.

El negocio aéreo es especialmente atractivo para el revenue Management por su alta perecibilidad, bajas barreras de entrada, oferta limitada en el corto plazo, distribución electrónica global y alta visibilidad de los precios.

Los beneficios otorgados por el R.M para el negocio, suelen ser claves para el éxito comercial ya que focaliza a la compañía en el crecimiento de los ingresos por sobre la reducción de los costos. “Las empresas que han aplicado técnicas de *Revenue Management* han visto crecer sus ingresos típicamente entre 3% y 7% y sus utilidades entre 50% y 100%”⁴.

El proceso de revenue Management se enmarca dentro del proceso de comercialización. Como se observa en la figura siguiente, este consta de 5 partes:

- Definición del producto: Consiste en definir la oferta en términos de servicios, políticas comerciales, estructura y restricciones tarifarias.
- Promoción: Publicidad, dar a conocer los productos, posicionamiento de marca.
- Distribución: Publicación de itinerarios, precios, venta en los diferentes canales disponibles.
- Maximización de ingresos: Maximización de ingreso neto al corto plazo sujeto a las restricciones de flota, itinerarios etc.
- Diagnóstico: Análisis y seguimiento para generar acciones de mejora.

⁴ Robert Cross. *Revenue Management: Hardcore Tactics for Market Domination*, 1997.



Figura 4: Proceso de comercialización [11]

En términos generales en LATAM este ciclo de gestión táctica de *Revenue Management* es concebido en 4 pasos principales:

1. Análisis e influencia Sistema (PROS): El sistema PROS es influenciado para que pueda realizar los pronósticos con la información más completa posible.
2. Pronostico demanda futura: PROS realiza el pronóstico de demanda basado en la información histórica de cada mercado.
3. Segmentación de la demanda: La demanda es segmentada de acuerdo a características específicas, usualmente anticipación de compra y estadía. Este procedimiento permite conocer mejor a los clientes para así entregarles el producto más adecuado a su disponibilidad de pago.
4. Asignación de oferta/disponibilidad para cada segmento de la demanda: El pronóstico de demanda y posterior optimización en PROS asigna la disponibilidad en cada mercado.
5. Establecer estructuras tarifarias: Este proceso permite tener una estructura de precios adecuado a cada segmento de demanda, asignando tarifas en base a anticipación de compra y estadía. Esta tarea corresponde a una gestión de mediano plazo.

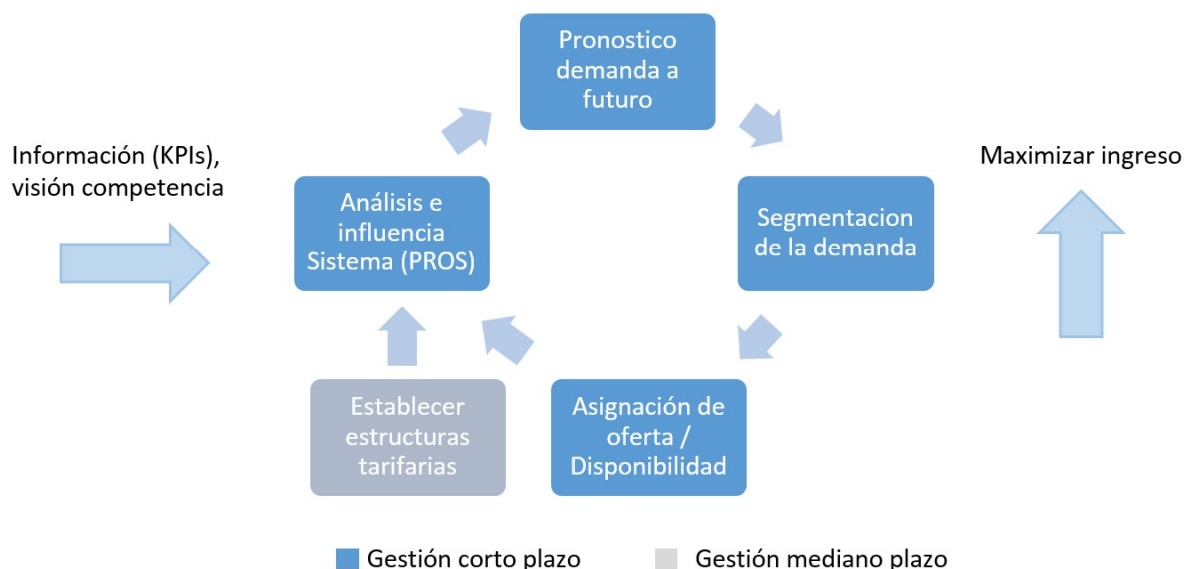


Figura 5: Modelo de trabajo actual táctico RM [11]

Estos 4 procesos tácticos se ven influenciados por el establecimiento de las estructuras tarifarias y todo el proceso concebido para maximizar el ingreso es alimentado con información. Es en esta última parte donde existe una brecha importante dado que no existen definiciones únicas, cada analista gasta mucho tiempo levantando información relevante y el correcto relacionamiento de los equipos de RM con otras áreas del negocio genera incoherencias en la información con la que otros equipos gestionan, es decir no existe una misma visión de la realidad.

Este proyecto se centra en poder unificar la información visualizada por los distintos equipos de Revenue Management para mejorar su coordinación y en el futuro, apuntar a un proceso que permita la conversación de distintas áreas involucradas en el proceso comercial.

Información (KPIs) del mercado: Se refiere a las principales características de la información del mercado que se esté gestionando. Estos datos pueden ser, por ejemplo, el market share, distribución de mercados de otras aerolíneas, la evolución de las reservas, etc. Esta información actualmente se obtiene de diversas fuentes. Posteriormente se aplican reglas de forma adhoc dependiendo de cada analista y su mercado. Sin embargo, no existe un sistema unificado, lo que conlleva un riesgo de errores en las consultas pues las reglas quedan a criterio de cada analista.,

2.2 Problemática de definiciones

Los problemas o desafíos que es necesario superar para lograr tener un sistema unificado de información son los siguientes

- **Definición de KPIS:** Coordinación con las áreas de negocios referentes a qué indicadores o KPIs son los relevantes para la correcta coordinación de los equipos y necesarios para activar las palancas comerciales.
- **Dar forma al proceso comercial:** Proceso recurrente que este proyecto pretender apoyar y dar forma, es el proceso por el cual las áreas se coordinan y logran llevar a cabo acciones concretas en el mercado observado dados los *findings* en la información.
- **Sistemas y fuentes de información:** Existen numerosas fuentes de información y sistemas dentro de la compañía, por lo tanto, no existe una única versión de los indicadores los que pueden inducir a errores posteriores.

2.3 Sistemas de información de LATAM

La gestión de la aerolínea hoy se basa en 5 fuentes de información estructurales que provienen de sistemas comerciales o en ciertos productos que se adquieren a proveedores. Estos son:

GPNR⁵. Base de reservas de la aerolínea, se estructura a través de los envíos diarios de archivos de los sistemas GDS/HOST, estos contienen la información con el detalle de las reservas activas en el sistema junto con las modificaciones que han tenido en el tiempo, estos archivos se denominan CDD.

Los archivos son enviados diariamente y procesados para contar con una base de datos actualizada diariamente. Estos archivos son un reflejo del sistema de reservas, y dado que es una observación al día, al momento de construir la base de información esta tiene un desfase con el tiempo real.

Información de Inventario (IFR): Por sus siglas Input Flight Report, corresponde a la fuente de datos de inventario de la aerolínea. Se construye en base al output que va desde el sistema de gestión de pasajeros (HOST) hacia el sistema de optimización (PROS). Contiene toda la información de capacidad, disponibilidad y utilización de los vuelos.

⁵ GPNR: Gestión PNR (Siglas de "Passenger Number Record"). El PNR corresponde al código de una reserva.

MIDT⁶: Por sus siglas en inglés *Market Information Data Transfer*, contiene la información de reservas provenientes de canales indirectos de la industria aérea, principalmente agencias de viajes. Es posible diferenciar la información a nivel de aerolínea y contiene los detalles del itinerario de cada reserva.

Esta información es muy utilizada para la comparación de reservas y pasajeros con respecto a otras aerolíneas en diferentes mercados y cálculo de participación de mercado.

JURO/TCN. Es la fuente de ingreso por concepto de tickets, posee la información en detalle del pasajero y el itinerario. Esta fuente tiene exclusivamente información de la compañía LATAM. Algunos de los indicadores más usados obtenidos desde esta fuente son el ingreso total y tarifas medias. Esta fuente es construida principalmente luego del traspaso de los archivos TCN/ISR⁷ que dispone un proveedor industria ATPCO⁸. La información contenida en esta fuente corresponde a un ingreso potencial generado por la venta, ya que en la contabilidad de la compañía el ingreso es devengado cuando el cupón es volado por el pasajero.

Cupón Volado. Es la base datos de información de tickets que ya fueron usados, es decir que corresponden a vuelos pasados. Esta fuente de información contiene ingresos de vuelos. A diferencia de Juro, acá se entrega la información contable por lo que la información de ingresos es de mejor calidad.

⁶ MIDT: *Market information data transfer*

⁷ TCN/ISR: *Transmission Control Number / Industry Sales Record*

⁸ ATPCO *Airline Tariff Publishing Company*

3. METODOLOGÍA

La metodología consta de cuatro partes alineadas para cumplir los objetivos planteados en este proyecto. Esta metodología se resume en:

- Entendimiento de la problemática y requerimientos principales del sistema
 - Dinámica del negocio general
 - Requerimientos principales
- Diseño lógico de métricas y atributos
 - Levantamiento de métricas y atributos generales
 - Métricas específicas de proyección de demanda
 - Selección de datos de muestra
 - Elección de modelos de proyecciones
- Diseño del modelo de fuentes de información y procesos
 - Identificación de fuentes de información
 - Diseño de procesos
 - Elección de herramientas
- Visualización
 - Diseño de visualización
 - Elección de la plataforma visual
 - Medición de resultados, métricas específicas

A continuación, un detalle de cada una de las secciones.

3.1 Entendimiento de la problemática y requerimientos principales del sistema

3.1.1 Dinámica del negocio general.

Debido a que este proyecto busca mejorar los resultados de las decisiones asociadas a procesos de coordinación entre los diversos equipos comerciales, es necesario primero conocer la dinámica del negocio que rodea a este proceso, especialmente desde el punto de vista de la información considerada para la toma de decisiones y la teoría detrás de ello.

Durante esta etapa se identifican las dinámicas del negocio y se genera un entendimiento de las brechas que pueda tener el proceso.

3.1.2 Requerimientos principales

Una vez finalizada la etapa anterior, es necesario levantar los requerimientos para el sistema de información, tanto en su contenido (qué métricas, indicadores, atributos y KPIs) como cuál será la manera apropiada para la navegación, periodicidad etc.

Estos requerimientos funcionales y de sistema serán posteriormente profundizados en detalle.

3.2 Diseño lógico de métricas y atributos

3.2.1. Levantamiento de métricas y atributos generales

Teniendo las especificaciones globales del sistema de información es necesario hacer un levantamiento de todas las métricas y atributos que serán utilizados y cuál es la fórmula de cálculo de cada una.

Dada la complejidad del negocio aeronáutico, esta labor no es trivial dado que existen muchas formas de ver un mismo indicador, dependiendo del uso que tenga en el área de negocio específica. Adicionalmente, debe existir un consenso entre los diversos equipos para lograr una misma visión.

3.2.2. Métricas específicas de proyección de demanda.

Una de las peculiaridades del negocio aéreo es que la demanda acumulada para un mercado va en aumento a través del tiempo hasta la fecha en que los vuelos de dicho mercado operan. La demanda se va acumulando conforme se van realizando las compras. Estas adquisiciones anticipadas o intenciones de viaje se van reflejando en los sistemas transaccionales como reservas. Son llevadas a observaciones temporales para poder registrar la evolución de la cantidad de reservas en los mercados. Por ejemplo, 100 reservas observadas el día 15 de agosto para un vuelo del día 20 de agosto.

En la literatura existen diversos *papers* con modelos propuestos que permiten predecir demanda utilizando información de reservas aéreas. Por ello, es necesario estudiar a profundidad cuáles son los modelos adecuados para este caso. Se debe recalcar que estas métricas son de apoyo y no pretenden resolver la problemática de la proyección de pasajeros, esta complejidad se escapa del alcance de este trabajo.

3.2.2.1 Selección de datos de muestra

Para los modelos de proyecciones es necesario contar con información histórica de reservas. Para ello, se debe seleccionar las bases y fuentes de información que

son adecuadas para este trabajo, la información que contienen, el tipo de reservas etc.

La selección de datos de venta es clave, ya que afecta directamente el desempeño de la herramienta a elegir. En general se espera, de acuerdo a lo estudiado, que el desarrollo de esta sección del sistema de información se realice con bases de datos de gran volumen y en permanente actualización.

3.2.2.2. Elección de modelos de proyecciones

Posterior a la selección de datos para la muestra, se aplicarán modelos de proyecciones. Éstos serán sometidos a mediciones para evaluar cuáles serán implementados.

Es importante recalcar que el desarrollo del proyecto considera una implementación, por lo que dentro de los requerimientos para los modelos a elegir es que sean programables y aplicables a grandes volúmenes de datos.

Por otro lado, muchas veces el contar con modelos complejos no asegura necesariamente una buena calidad en la proyección. Sumado a esto la dificultad que añade un modelo complejo a la gestión del cambio, puesto que convencer a los diferentes actores se hace más difícil conforme el modelo se complejiza.

Además, es relevante notar que muchas veces el contar con modelos complejos no asegura buena calidad en la proyección además que la dificultad para

3.3 Diseño del modelo de fuentes de información y procesos

3.3.1 Identificación fuentes de información

Luego del levantamiento de requerimientos y coordinación de las áreas comerciales, es necesario establecer cuáles son las fuentes de información que darán soporte a las métricas y atributos mapeados.

Es importante identificar en las fuentes de información sus elementos particulares, cómo pueden hacer que la data contenida no sea comparable, capturar estas diferencias y poder explicarlas.

3.3.2 Diseño de procesos

Estableciendo la metodología de cálculo y entendiendo las particularidades de las fuentes es necesario diseñar los procesos por los cuales estos datos serán transformados en información, es decir en indicadores y atributos.

Estos procesos de transformación de datos, conocidos como ETL, permiten dar forma a la data para así entregar información de acuerdo a las reglas de cálculo y lógicas establecidas anteriormente.

Esta es una tarea que requiere de habilidades mixtas de uso de herramientas tecnológicas y una comprensión de negocio para lograr resultados exitosos.

3.3.3 Elección de herramientas

Para la realización del piloto es necesario elegir herramientas que permitan una rápida implementación, que sea escalable y capaz de manejar un gran volumen de datos.

Los criterios que se deben tener en consideración para la elección son los siguientes:

- Herramientas ya existentes: Se debe utilizar herramientas que ya se encuentran presentes en la compañía, en la medida de lo posible, para así reducir los costos de gestión del cambio al momento de la implementación.
- Herramientas nuevas, escalables y open source: Para el caso en que las herramientas requeridas no se encuentren en uso dentro de la compañía, es recomendable elegir las que permitan una rápida escalabilidad, usabilidad y que sean open source, para una mejor distribución dentro de los futuros usuarios.

3.4 Visualización

3.4.1 Diseño de visualización

La visualización es muy relevante a la hora de construir una herramienta que será utilizada por diversos usuarios y que tendrá por objetivo la coordinación de equipos, los cuales internamente utilizan diferentes aperturas de los indicadores incluidos en el sistema de información.

Es por ello que se opta por una forma de visualización ampliamente extendida por la compañía y que es familiar para todos los usuarios, como es Microsoft Excel.

La capacidad y flexibilidad que otorga esta herramienta, sumado al conocimiento que tienen los usuarios de ésta la hacen idónea para diseñar en ella una visualización que pueda ser customizada por los usuarios, haciendo más fácil la

adopción de la herramienta por parte de ellos y aportando hacia la gestión del cambio.

3.4.2 Elección de plataforma visual

Como se mencionó anteriormente, la construcción de la visualización es muy relevante para este sistema de información, debido a que la herramienta tiene por objetivo servir de apoyo a la coordinación de los equipos.

Los datos deben estar agrupados y ser mostrados de manera simple pero a la vez poderosa, con capacidad para cambiar la visualización de forma rápida dependiendo de las necesidades del equipo que está realizando los análisis.

Esto representa un desafío técnico, pues las plataformas usadas comúnmente como Microsoft Excel no pueden resolver el manejo del extenso volumen de información, crítico en para el éxito de la herramienta.

Es recomendable la utilización de Microsoft Excel como visualizador debido a que se encuentra ampliamente inserto dentro de la organización, sin embargo, es necesario contar con un complemento a este para el procesamiento de datos y así conservar todas las capacidades de Excel sin perder el desempeño de la herramienta.

3.4.3. Medición de resultados, métricas específicas de proyección

Una vez estudiados los modelos de proyección de demanda, se debe evaluar cuál o cuáles ofrecen los mejores resultados, utilizando los datos históricos para los distintos mercados en los que opera la compañía.

Se utilizarán indicadores estadísticos para medir el desempeño de los modelos:

- Estándares de la compañía: Los resultados obtenidos por los modelos de proyección deberán estar dentro de un cierto margen definido por el propio contexto del mercado en donde se está proyectando. Los resultados no podrán ser menores que las reservas confirmadas y mayores que la capacidad total.
- Error estadístico: Se revisará el error estadístico de las observaciones para poder evaluar la precisión del modelo.

4. MARCO CONCEPTUAL

Para el cumplimiento de los objetivos de este trabajo es necesario revisar los conceptos del negocio aéreo y los principales indicadores de gestión. También es importante que el lector se familiarice con los conceptos de la operación comercial de una aerolínea.

4.1 Conceptos comerciales y métricas del negocio de transporte

4.1.1 Gestión comercial y sistemas de distribución

La gestión comercial de una aerolínea es una larga cadena de decisiones que va desde las definiciones estratégicas de flota (cantidad y tipo de aviones), países y aeropuertos de operación, itinerarios, entre otras, hasta problemáticas de negocio de corto plazo que denominaremos decisión de la operación comercial. Es en este último tipo de decisiones en el que se aboca el presente trabajo.

Como ejemplo de lo anterior, se observa en la Figura 6 los plazos que se estipulan para cada tipo de decisión. Las decisiones de flota pueden tomar hasta más de un año, mientras que el horizonte de tiempo de las decisiones de Revenue Management no excede los 3 meses.

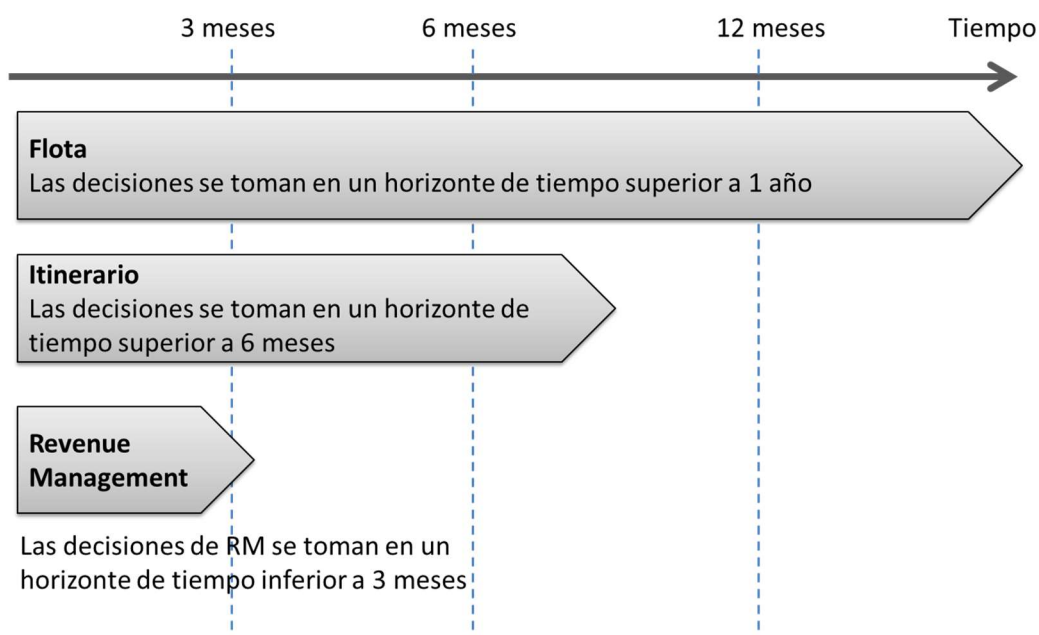


Figura 6: Decisiones gestión comercial

La operación comercial gira en torno a los GDS (*Global distribution system*) que corresponden a sistemas que permiten a las aerolíneas ofrecer vuelos en todas partes del mundo, y que conectados a los sistemas de inventario permiten también especificar sofisticadas reglas.

Los GDS nacieron a mitad del siglo pasado, y en la actualidad representan un costo muy relevante en la operación de una aerolínea. Un porcentaje del costo de cada pasaje equivale a costo del GDS.

Hoy en día los GDS con mayor participación de mercado son Sabre (US) y Amadeus (Francia). Ambos tienen alrededor del 60% del mercado de las aerolíneas en el mundo. En Latam se utiliza Sabre para vuelos LAN y Amadeus para vuelos TAM.

El 5 de mayo de 2015 LATAM Airlines informo que había elegido a Sabre como el principal cooperador tecnológico y por lo tanto se dio inicio a un proceso de cambio desde los sistemas de Amadeus en TAM para incorporarse a Sabre. Esto conlleva un proceso de cambio tecnológico, de sistemas, procesos y procedimientos que se espera culmine en octubre de 2017.

4.1.2 Conceptos geográficos y de vuelo

La conceptualización geográfica es bastante estándar en la industria aérea y es relevante para el entendimiento de las dinámicas del día a día. Dentro de éstos conceptos podemos encontrar los siguientes:

Aeropuertos. Corresponde a la unidad básica geográfica en el negocio aéreo, nos indica el punto de partida o llegada de un avión o pasajero. Por ejemplo, el aeropuerto de Santiago de Chile, "Arturo Merino Benítez", es denominado **SCL** según la nomenclatura estándar internacional (Normado por la IATA)

IATA: "International Air Transport Association". Corresponde a una organización a la cual pertenecen más de 260 aerolíneas, es una entidad regulatoria que vela por la correcta convivencia entre las aerolíneas y el cumplimiento de estándares.

Tramo. Trayecto de un avión entre dos **aeropuertos** sin paradas intermedias. Por ejemplo, en el caso que el avión vuele desde el aeropuerto de Santiago, Chile al de Lima, Perú, la correcta denominación sería tramo: **SCLLIM** (Secuencia de códigos de los dos aeropuertos según denominación internacional ordenados de inicio a fin del viaje).

Vuelo. Corresponde a la denominación numérica que permite identificar el itinerario de un avión en específico. En el estándar internacional el vuelo corresponde a un código alfanumérico de dos letras (denominado carrier, perteneciente a la aerolínea que lo comercializa) seguidos de un número de hasta 4 dígitos. Por ejemplo, para el caso

descrito previamente tenemos el vuelo **LA 530** (Comercializado por LAN Airlines número de vuelo 530), este vuelo realiza el tramo **SCLLIM**. Es importante destacar que si se considera la fecha en que este vuelo está programado, esta sería una combinación única (por convención internacional), es decir, no pueden existir dos vuelos con la misma nomenclatura que operen el mismo día.

Segmento. Es la denominación para indicar uno o más **tramos** que realiza un pasajero en un mismo vuelo, su denominación corresponde a la secuencia de aeropuertos de inicio del primer tramo y del último tramo (ordenados secuencialmente por fecha y hora).

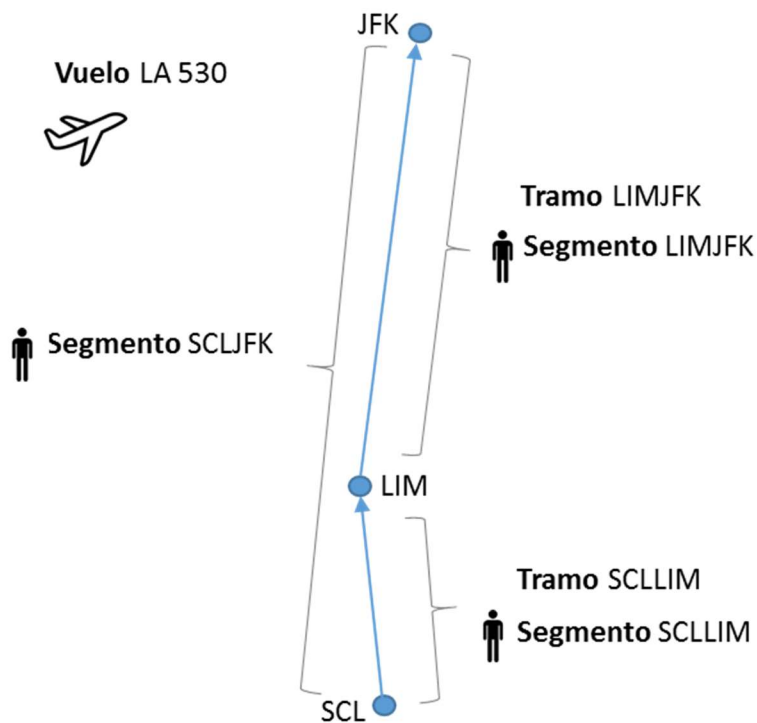


Figura 7: Ejemplo definición segmentos y tramos

En el diagrama se observa que desde el punto de vista del pasajero este pudo realizar el segmento Santiago-Lima que es equivalente al tramo SCLLIM que realiza el vuelo LA 530. Así también, otro pasajero distinto pudo realizar el tramo Lima-Nueva York a través del vuelo 530 y un tercer pasajero realizar el recorrido completo SCL-JFK. En este último caso no existe el tramo SCLJFK, esto porque el vuelo 530 realiza una parada en Lima. El concepto que engloba la intención de viaje de un pasajero es el Origen Destino

Origen-Destino. (Abrev. OD). Se denomina a la secuencia de los códigos de las dos ciudades correspondientes al origen y al destino del pasajero resultantes del modelamiento realizado de la intención de viaje a través del itinerario de la reserva, esto independiente de las conexiones que ese pasajero tenga y/o líneas aéreas involucradas dentro de su viaje.

Actualmente el modelamiento industria del OD se realiza de manera estadística, más específicamente se definen patrones asociados a los tiempos de conexión entre los diferentes vuelos. Durante el desarrollo del proyecto se ahondará en mayor detalle sobre como este concepto es concebido particularmente en LATAM Airlines.

Adicionalmente, se puede hablar de agrupación de ODs, y estos simplemente apunta a pasajeros que contienen orígenes y destinos agrupados sobre alguna jerarquía geográfica. Por ejemplo: la agrupación OD: CL-US corresponde a todos los orígenes destino que geográficamente pertenecen a estos 2 países.

Para este trabajo se asume los mercados proyectos son independientes entre sí.

Routing. Secuencia de números de vuelo y segmentos que un pasajero utiliza para moverse entre un origen y un destino. Por ejemplo: LA630 SCL-LIM-LAX AA1010 LAX-SFO. Aquí: LA630 SCL-LIM-LAX corresponde al *routing* “online” (parte solo operada por holding LATAM), donde LA630 es el vuelo n° de vuelo LAN y SCL-LIM-LAX son los 2 tramos que contienen el segmento SCL-LAX (SCL-LIM, LIM-LAX), el segundo segmento es el AA1010 LAX-SFO, es el *routing* “offline”, corresponde a un vuelo de American Airlines (AA), y va de LAX a SFO (Los Angeles, USA – San Francisco, USA).

4.1.3 Conceptos relacionados a la venta y canales

Reserva. Uno de los principales elementos para la trazabilidad del transporte de pasajeros. Este documento o registro único contiene todos los detalles del servicio a prestar, incluyendo los segmentos, vuelos y fechas respectivas de viaje, además de contener instrucciones para servicios especiales, comidas, mascotas, asiento, entre otros. Para efectos de este trabajo se abordarán en exclusiva los elementos relacionados al itinerario.

Las reservas pueden ser realizadas para personas individuales o también para grupos, en este último caso corresponden a 10 o más personas que realizan una reserva conjuntamente.

Ticket. Es el documento registra la compra de un pasaje aéreo previamente reservado, es el elemento que contiene el itinerario, las condiciones y tarifa con las que un pasajero adquirió el servicio.

POS (*Point of Sale*). País de venta del ticket aéreo, se señala con dos letras. Por ejemplo, CL corresponde a Chile.

Tarifa. Es el precio que debe pagar el pasajero por un determinado producto. Las aerolíneas publican sus tarifas diferenciadas por las condiciones del viaje y regulaciones y es una de las formas de capturar una mayor disposición a pagar y segmentar la demanda. Algunos de los principales elementos de diferenciación de las tarifas son:

- **Anticipación:** Diferencia entre la fecha de compra y la fecha del vuelo.
- **Max/min estadía:** Cantidad de días entre la fecha de ida y la fecha de regreso.
- **Restricciones para días específicos (feriados, fines de semana, etc.):** Se aplican restricciones para fechas especiales en donde podría haber una mayor disposición a pagar por parte del pasajero.
- **Stopovers:** Estadías en puntos intermedios del viaje.
- **Restricciones de vuelos:** En algunas ocasiones las tarifas aplican solo para ciertos vuelos, de tal forma de segmentar la demanda por el producto ofrecido. Por ejemplo, el vuelo directo vs el vuelo en conexión.

Dadas estas restricciones es posible que algunos vuelos no estén disponibles a la venta aun cuando todavía quede espacio libre en ellos.

Segmentación de demanda. Proceso mediante el cual se clasifican los pasajeros por algunos parámetros que caracterizan su viaje. Así, es posible entregar la tarifa más adecuada al pasajero y su intención de viaje de esa ocasión. La segmentación de demanda en LATAM se describe como:

- **Pasajero turista:** El más común, se caracteriza por estadías de 1 semana o más, incluyendo fines de semana, alta anticipación, con equipaje y viajando con una persona o más.
- **Pasajero étnico:** Se refiere a los pasajeros que viajan por motivos familiares. Se caracteriza por largas estadías, alta anticipación y viajes sobre todo en fechas específicas (Navidad, Año nuevo, otras fiestas)
- **Pasajero corporativo o de negocios:** El pasajero con mayor yield, se caracteriza por su baja anticipación, cortas estadías, sin noche de viernes o noche de sábado y usualmente viaja solo y con maleta de mano. Adicionalmente este tipo de pasajeros se interesa por horarios específicos de vuelo (por ejemplo, vuelos en la mañana para la ida y en la tarde para la vuelta, o bien vuelos nocturnos para el caso de viajes de más de 8 horas de duración)

Clase tarifaria. Es la representación interna de la empresa que clasifica la tarifa dentro de un rango y es dependiente del mercado para las restricciones previamente descritas. Las restricciones tarifarias de las clases son las que permiten adecuar la tarifa al segmento de demanda del pasajero que adquiere un boleto.

Cabina. Es el espacio físico o lógico dentro del avión donde se sientan los pasajeros de acuerdo a las diferentes clases tarifarias. Ejemplo: Cabina Económica alberga las clases tarifarias más bajas.

Canal Venta directa. Se habla de venta directa cuando el ticket o reserva aéreo ha sido adquirido en una oficina o call center de la compañía. Esto también aplica para la compra dentro del sitio web de Latam. La persona, natural o jurídica que compra el ticket es la misma que lo va a usar, no hay intermediarios.

La venta directa se realiza a través de los call centers, *counters* de Aeropuerto (*ATO: Airport Ticket Office*), oficinas de venta dentro de la ciudad (*CTO: City Ticket Office*) y a través de internet (*Lan.com o tam.com*)

Canal Venta indirecta. Se habla de venta indirecta cuando el pasajero compra el ticket a un intermediario, es decir, una entidad ajena a la compañía, en este caso el que compra el ticket es distinto de quien lo va a usar. Esta venta en la mayoría de los casos genera una comisión.

En este grupo podemos encontrar principalmente a las agencias de viaje físicas y a las OTAS (*Online travel agencies*), como Booking.com, Despegar.com, etc.

4.1.4 Indicadores comerciales y de rentabilidad

Tarifa media. Corresponde a la media de ingreso por pasajero, esto calculado al mismo nivel de aperturas se describe de la siguiente manera

$$TM = \frac{\text{Tarifa}}{\text{Pasajero}}$$

ASK. (*Available Seat Kilometer*) Es una medida de capacidad u oferta de transporte aéreo y se calcula como el total de asientos ofrecidos multiplicados por la distancia volada. Es un indicador de la oferta de la línea aérea. Los ASK's pueden calcularse para un tramo, un vuelo determinado, una aerolínea y/o la industria aérea en su totalidad.

$$ASK = \text{Número de asientos} * \text{kilómetros volados}$$

$$\text{Asiento} * 1 \text{ Kilómetro} = 1 \text{ ASK}$$

RPK. (*Revenue Passenger Kilometers*) Es el número de pasajeros multiplicados por el número de kilómetros volados. Es una medida de la demanda que tiene la línea aérea. Los RPKs pueden medirse para un tramo, un segmento o par de puntos, un vuelo determinado, para una aerolínea y/o para la industria aérea en su totalidad

$$\text{RPK} = \text{Cantidad de pasajeros pagados} * \text{Kilómetros volados}$$

$$\text{Pasajero} * 1 \text{ km} = 1 \text{ RPK}$$

FO. (Factor de ocupación) Es un índice de cuan copados están los vuelos y sus tramos. Cada ruta tiene un factor de ocupación mínimo para que un vuelo sea rentable a un determinado precio.

El factor de ocupación mínimo varía según las rutas, ya que depende de las tarifas, de la carga transportada, del costo de ese vuelo para la línea aérea, etc.

El FO se puede calcular para un tramo, vuelo, una ruta, y/o para la línea aérea como un todo:

$$FO = \frac{\text{Pasajeros por kilómetro volado}}{\text{Asientos por kilómetro volado}} = \frac{RPK}{ASK}$$

BELF (*Break Even Load Factor*): Es el factor de ocupación mínimo tal que la ruta iguala costos e ingresos para comenzar a ser rentable.

Yield. Es un indicador o medida de rentabilidad asociada normalizada que permite la comparación unificada de vuelos para diferentes distancias, es el ingreso de la aerolínea por pasajero-kilometro

Dado lo anterior, para una línea aérea son más interesantes las rutas de alto yield. El pasajero de negocios es de más alto Yield que el étnico o de turismo.

El Yield puede calcularse para un vuelo, para una ruta y/o para la línea aérea como un todo.

$$\text{Yield} = \frac{\text{Ingreso}}{RPK} * 100$$

RASK (*Revenue per Available Seat Kilometer*) Es una medida de ingreso por unidad de oferta de transporte aéreo. Se calcula como la división entre el ingreso y los asientos disponibles multiplicados por la distancia volada. Este KPI es un indicador de eficiencia en la generación de ingresos.

$$RASK = \frac{\text{Ingreso}}{ASK} = \text{Yield} * FO$$

CASK (*Cost per Available Seat Kilometer*) Es una medida del costo de la oferta de transporte aéreo y se calcula como la división entre el costo total y los asientos disponibles multiplicados por la distancia volada.

$$CASK = \frac{\text{Costo}}{ASK} * 100$$

4.2 Modelos y métodos cuantitativos para el pronóstico de demanda.

Como parte de los objetivos del trabajo se realizará un estudio e implementación que permita desplegar una proyección de demanda de pasajeros, al mismo nivel de granularidad que los analistas de RM (*Revenue Management*) proyectan y gestionan. Esto sólo aplica para entregar un nivel mayor de comprensión de los diversos mercados y no como reemplazo o intervención de la proyección que realizada por los analistas.

4.2.1 Modelos predictivos

Los modelos predictivos son técnicas estadísticas que permiten estimar un comportamiento futuro. Permiten extraer patrones desde los datos históricos para identificar riesgos y oportunidades. Estos modelos se componen de predictores que explican algún evento o situación que se desea inferir.

Las técnicas de análisis predictivo se separan en:

Técnicas de regresión: Apuntan a entregar un modelo matemático representativo de la interacción de las variables en consideración.

Técnicas ‘Machine Learning’: Apuntan a emular un modelo cognitivo y adquirir aprendizaje a través del entrenamiento. Estos modelos funcionan como una ‘caja negra’ donde las variables explicativas no son explícitas.

Para efectos de este proyecto, se trabajará con técnicas de regresión, las cuales entregan una representación matemática del comportamiento de los elementos y permite ofrecer una idea más clara del comportamiento asociado a la proyección, generando un aprendizaje deseado en los analistas.

4.2.2 Modelo básico económico de oferta y demanda en una aerolínea

Oferta

El concepto de oferta en una aerolínea corresponde a la capacidad máxima de asientos disponibles a la venta, independiente del origen o destino final de los potenciales pasajeros que la ocupen. Este valor se mide de acuerdo a lo descrito previamente como ASK.

El producto en una aerolínea se define por las variables origen, destino, vuelo, fecha. También puede ser ruta, vuelo, fecha o bien definido a nivel de mercado. Por otra parte, el inventario es definido por la combinación entre los productos disponibles por los asientos que tengo para cada una de esas rutas. De esta forma, se deduce que, debido a que tanto los productos como los asientos son limitados, entonces la oferta en una aerolínea es finita.

Demanda

El concepto de demanda en la teoría económica es descrito como la cantidad de un bien o servicio que los consumidores (interesados) están dispuestos a pagar, a un precio y en un periodo determinado. La demanda tiene una relación inversa con el precio de un producto, es decir manteniendo todos los otros factores iguales una baja de precio tiende a aumentar la demanda.

Para el caso de la aerolínea, la demanda viene dada por los pasajeros que pagan su pasaje en una determinada ruta, vuelo, fecha.

Elasticidad precio

La relación entre el precio y la demanda se conceptualiza junto con la oferta como la elasticidad precio demanda:

$$\frac{\partial Q}{\partial P} \times \frac{P}{Q}$$

Existen diversas tipificaciones de la demanda que apuntan a armar grupos más reducidos para estudios particulares, algunas de estas separaciones son demanda externa, interna, latente, irregular, etc. Dentro del mundo del transporte una de las conceptualizaciones más utilizadas es la demanda *constrained* y *unconstrained*.

Demanda *constrained*: Demanda acotada a las condiciones y restricciones de la oferta, flujo de caja, capacidad disponible, entre otros.

Demanda *unconstrained*: Estimación de la demanda que habría llegado realmente si es que no se tuvieran las restricciones anteriores.

Se utilizará la capacidad agrupada al nivel de la proyección como la cota máxima para la estimación de demanda de pasajeros. En otras palabras, será imposible tener mayor demanda *constrained* que la capacidad máxima disponible para transportarlos.

Por otra parte, la cota inferior para la proyección son las reservas de pasajeros que ya poseen reserva. Es decir, no se podrá tener una demanda menor que las reservas que ya se encuentran registradas en ese momento en el sistema.

En conclusión, la proyección de demanda *constrained* está acotada entre la cantidad de pasajeros por volar ya reservados y la capacidad máxima de transporte.

$$\sum Pa \text{ Reserva} < \sum \text{Demanda Constrained} \leq \sum \text{Capacidad Total}$$

4.2.3 Series de Tiempo

Una serie de tiempo es una forma estructurada de representación de datos, como una curva que evoluciona en el tiempo. Representan una relación entre un valor asociado a la ocurrencia de determinado fenómeno y una fecha determinada, pudiendo así observar la variación de las realizaciones en el tiempo.

Por ejemplo, al analizar la demanda de pasajeros para un vuelo, la variable a observar es la cantidad de pasajeros para un determinado origen, agrupados por día, semana, mes o año.

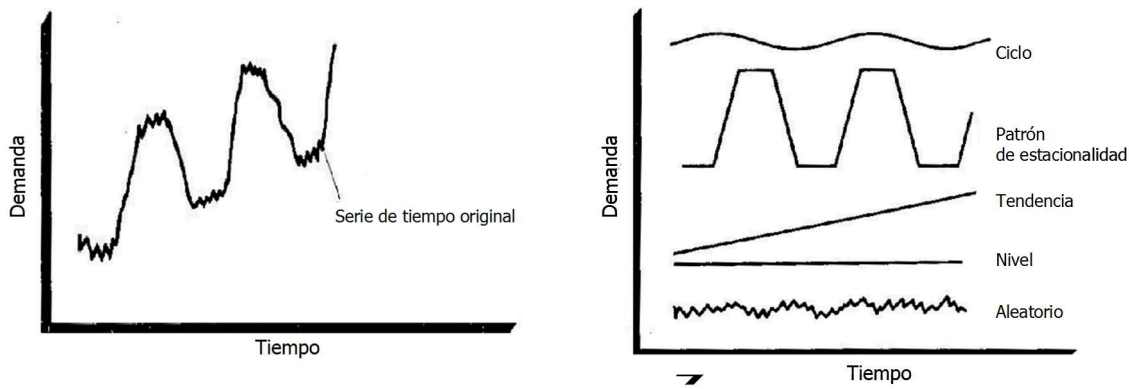


Figura 8: Ejemplo de series de tiempo

El pronóstico de series de tiempo implica una extensión de las variables históricas hacia el futuro, donde no se tienen variables disponibles. Se caracterizan en dos variables estructurales principales:

- **Periodo:** Representa el nivel de agregación. Pueden ser días, semanas, meses, etc. Para el caso de la aerolínea la menor granularidad trabajada es a nivel diario.
- **Horizonte:** Representa la cantidad de periodos a pronosticar.

Como se observa en la Figura 8, una serie de tiempo puede ser descompuesta en 5 elementos que representan niveles básicos de variación. Todos estos, actuando al unísono, son los que contribuyen a los cambios observados en el tiempo.

Se asume que existe una relación multiplicativa entre las componentes, por lo que un cambio en el valor de la serie puede ser atribuido a la acción de las componentes en conjunto.

Estos elementos son:

Ciclo: Frecuentemente las series de tiempo presentan secuencias alternas de puntos arriba y debajo de la línea de tendencia con una duración prolongada en el tiempo. La variación se mantiene cuando se ha eliminado las variaciones o tendencia estacional e irregular. Por ejemplo, los ciclos económicos que presentan periodos recurrentes de prosperidad, recesión, depresión y recuperación.

Patrón de estacionalidad: Componente de la serie de tiempo que representa variabilidad en los datos por influencia en las estaciones. Corresponde a los movimientos de la serie que son recurrentes, apareciendo regularmente en periodos de tiempo y niveles de intensidad similares. Por ejemplo, la venta de trajes de baño puede ser modelada como una serie de tiempo estacional, manifestando un periodo de alza durante los meses de verano.

Tendencia: En una serie de tiempo, la tendencia es la característica que muestra el patrón gradual y consistente de las variables de la serie, que considera las fuerzas externas y persistentes que afectan el crecimiento o reducción de la misma. Corresponde por lo general a factores de largo plazo que tienen efecto en los resultados. Por ejemplo, los cambios poblacionales.

Aleatorio o variación irregular: Efectos de corto plazo, impredecibles y no recurrentes en la serie de tiempo. Se pueden clasificar en las variaciones que son provocadas por acontecimientos especiales (por ejemplo terremotos, inundaciones, etc) y variaciones aleatorias, en las cuales no es posible identificar las causas de forma exacta.

Dado lo anterior, es posible describir una serie temporal como:

$$X_t = T_t + E_t + I_t$$

Donde T_t es la tendencia, E_t es la estacionalidad e I_t es la componente aleatoria.

Dependiendo de la estacionalidad de la serie, se puede clasificar en dos tipos:

- **Serie estacionaria:** este tipo de series tienen una estabilidad a lo largo del tiempo, que se puede identificar gráficamente al ver que los valores de la serie oscilan alrededor de la media y la variabilidad es constante.
- **Serie no estacionaria:** Cuando la variabilidad y tendencia cambian en el tiempo. Los cambios en la media indican una tendencia creciente o decreciente en el largo plazo.

La tendencia de una serie puede clasificarse en dos:

- **Tendencia lineal:** Indica un aumento o disminución de una variable a un ritmo constante. Frecuentemente se utilizan tendencias lineales como descriptor de las variaciones al largo plazo.
- **Tendencia no lineal:** Cuando la serie de tiempo tiene un comportamiento curvilíneo, entonces es no lineal. Frecuentemente se utilizan las tendencias polinomial, logarítmica, exponencial y potencial.

4.2.3.1. Métodos de pronóstico en series de tiempo

A continuación, se detallan los modelos más frecuentemente utilizados para las series de tiempo. La notación a utilizar será la siguiente:

D_t : Demanda observada en el periodo t

F_{t+1} : Pronóstico para el periodo t

A_t : Promedio calculado para el periodo t

$E_t = D_t - F_t$ Error de pronóstico

- a. **Promedio simple:** Atenúa los datos al obtener la media de un conjunto de observaciones históricas para tener el pronóstico del periodo siguiente. Es utilizado en series sin patrones de tendencia, estacionalidad o ciclo. Se calcula como:

$$F_{t+1} = \sum_{t=1}^n \frac{D_t}{n}$$

b. Promedio Móvil Simple

Otorga una mayor importancia al dato más reciente, pues al agregar una nueva observación elimina la más antigua. Entonces, dado un periodo de N observaciones, siempre se mantendrá esa cantidad de datos al momento de realizar el pronóstico siguiente.

$$A_t = \frac{D_t + D_{t+1} + \dots + D_{t-N+1}}{N}$$

$$F_{t+1} = A_t$$

c. Promedio móvil doble

Calcula un conjunto de promedios móviles y luego calcula un segundo promedio móvil a partir del primero. Se utiliza en pronóstico de series con tendencia lineal.

El primer promedio móvil es calculado como:

$$A_t = \frac{D_t + D_{t+1} + \dots + D_{t-N+1}}{N}$$

Y el segundo se calcula como:

$$A_{t'} = \frac{A_t + A_{t+1} + \dots + A_{t-N+1}}{N}$$

Luego se calcula la diferencia entre los dos promedios móviles

$$a_t = 2A_t - A_{t'}$$

Y se añade un factor adicional para el ajuste

$$b_t = 2 * \frac{A_t - A_{t'}}{n - 1}$$

Con ambos factores se obtiene una expresión para pronosticar m periodos en el futuro:

$$F_{t+m} = a_t b_t m$$

d. Promedio Móvil Ponderado

Este método permite asignar importancia o peso específico a cada dato de los promedios, siempre que la suma de las ponderaciones sea equivalente a 100%. Usualmente se aplica el mayor factor de ponderación al dato más reciente.

Este método es frecuentemente utilizado para patrones de demanda aleatoria o nivelada donde se desee eliminar el impacto de los elementos históricos irregulares, enfocándose en la demanda más reciente.

$$F_{t+1} = A_t = W_t * D_t + \dots + W_{t-N+1} * D_{t-N+1}$$

Donde se debe cumplir que

$$\sum_{i=t-N+1}^t W_i = 1$$

e. Alisamiento Exponencial Simple:

Realiza una atenuación de los valores de la serie de tiempo, obteniendo el promedio de forma exponencial. Así, se da más peso a las observaciones más recientes. El promedio se calcula como:

$$A_t = \alpha D_t + (1 - \alpha)A_{t-1}$$

$$F_{t+1} = A_t$$

f. Alisamiento Exponencial con Tendencia (Método de Brown):

Similar al método anterior, es utilizado para casos en los que se presenta una tendencia en la serie de tiempo. Esto permite mejorar los pronósticos ajustando el modelo del alisamiento exponencial a los desfases de la tendencia.

En este caso:

A_t Demanda real en el periodo t

T_t Tendencia suavizada exponencialmente en el periodo t

α Constante de alisamiento para la media

β Constante de alisamiento para la tendencia

$$A_t = \alpha D_t + (1 - \alpha)(A_{t-1} + T_{t-1}) \quad 0 < \alpha < 1$$

$$T_t = \beta(A_t - A_{t-1}) + (1 - \beta)T_{t-1} \quad 0 < \beta < 1$$

$$F_{t+k} = A_t + kT_t$$

g. Alisamiento Exponencial con Estacionalidad de Largo L (Método de Holt Winters):

Cuando existe una variación estacional, los modelos aplicados deben tener una componente de corrección para la corrección de la estacionalidad que existe en las observaciones.

En este caso,

A_t Demanda real en el periodo t

R_t Estacionalidad suavizada exponencialmente en el periodo t

α Constante de alisamiento para la media

γ Constante de alisamiento para la estacionalidad

$$A_t = \alpha \left(\frac{D_t}{R_{t-L}} \right) + (1 - \alpha)A_{t-1} \quad 0 < \alpha < 1$$

$$R_t = \gamma \left(\frac{D_t}{R_{t-L}} \right) + (1 - \gamma)R_{t-L} \quad 0 < \gamma < 1$$

$$F_{t+k} = A_t * R_{t-L+k}$$

h. Alisamiento Exponencial con Estacionalidad de Largo L y Tendencia:

Cuando existe una variación estacional y una variación de tendencia, se debe contar con componentes correctores para ambos efectos.

En este caso,

A_t Demanda real en el periodo t

T_t Tendencia suavizada exponencialmente en el periodo t

R_t Estacionalidad suavizada exponencialmente en el periodo t

α Constante de alisamiento para la media

β Constante de alisamiento para la tendencia

γ Constante de alisamiento para la estacionalidad

$$A_t = \alpha \left(\frac{D_t}{R_{t-L}} \right) + (1 - \alpha)(A_{t-1} + T_{t-1}) \quad 0 < \alpha < 1$$

$$R_t = \gamma \left(\frac{D_t}{A_t} \right) + (1 - \gamma)R_{t-L} \quad 0 < \gamma < 1$$

$$T_t = \beta(A_t - A_{t-1}) + (1 - \beta)T_{t-1} \quad 0 < \beta < 1$$

$$F_{t+k} = (A_t + T_t)R_{t-L+k}$$

Para el proyecto, se requiere calcular un pronóstico de demanda de pasajeros como atributo adicional a los KPIs definidos como relevantes por los equipos de negocio involucrados. Por ello, es necesario realizar pruebas con varios modelos y determinar el más adecuado.

Es posible medir la precisión de los modelos con diversas métricas, que asocian la calidad del modelo con el error.

4.2.3.2. Errores de Pronóstico:

a. Desviación media absoluta (MAD)

$$MAD = \sum_t \frac{|D_t - F_t|}{n}$$

Donde n representa el número de periodos para la inicialización.

b. MAD para pronósticos expost

$$MAD_t = \alpha |F_t - D_t| + (1 - \alpha)MAD_{t-1}$$

Este indicador de error permite identificar valores extremos con el siguiente criterio:

Si $|F_t - D_t| > 3,75 * MAD_t$ (tres desviaciones estándar), entonces existen sospechas de la presencia de un valor extremo.

Adicionalmente, en un enfoque aproximado, σ corresponde a $1,25 * MAD$ para una función normal.

c. Señal de Rastreo:

La señal de rastreo es una medida de desempeño que permite medir la desviación del pronóstico respecto a variaciones de la demanda.

Los límites aceptables de la señal de rastreo dependen del tamaño de la demanda pronosticada, sin embargo, un rango aceptable es que TS se encuentre entre $[-6,6]MAD$. Si la señal de rastreo no se encuentra en ese rango (confiabilidad menor al 97%), se debe revisar el modelo ya que está sesgado.

La señal de rastreo se calcula como:

$$TS = \frac{\text{Suma acumulada de la desviación del pronóstico}}{MAD}$$

O bien,

$$TS = \frac{\sum_{t=1}^n (A_t - F_t)}{\frac{1}{n} \sum_{t=1}^n |A_t - F_t|}$$

Adicionalmente, la señal de rastreo entrega información sobre el pronóstico y la demanda: Cuando la demanda real excede al pronóstico, el TS es positivo. Por el contrario, cuando el TS es negativo, entonces la demanda es menor al pronóstico.

4.2.3.3. Técnicas pickup

A diferencia de estimar demanda con absoluta incerteza, para las reservas aéreas el elemento a proyectar se va “llenando” con el paso del tiempo, dado que existe una anticipación al uso del servicio, en este caso los vuelos futuros. Es por esto que es necesario realizar una modificación a la aplicación tradicional de series de tiempo para incorporar la información adicional que se tiene al proyectar demanda en el negocio aéreo.

Una de las técnicas ampliamente usadas en RM que capturan la información de reservas parciales, es la técnica de pickups, este procedimiento consiste en separar la muestra en diferenciales de crecimiento en vez de números nominales y con esto predecir las

variaciones de las reservas futuras, de esta manera se utiliza la información de reservas en vuelos futuros en la proyección.

Formalmente se puede definir el elemento a proyectar como:

$$C_{i,j}^{Mult} = C_{i,j}/C_{i,j+1}, j = 0, 1, \dots, h.$$

Donde cada pick-up “C” es un factor de cambio en el caso de las reservas para un mes, semana o día de vuelo dependiendo de la frecuencia con que se requiera realizar la proyección. Lo importante de esta técnica es que el resultado de la proyección no es directamente el necesario, sino que son vectores multiplicativos. La variación multiplicativa de los *pickups* se utilizará dado que se han encontrado mejores resultados de esta manera (1).

4.3 Modelos de procesos y bases de datos

4.3.1 Modelamiento de procesos de negocio

Un modelo es la representación de la realidad. Particularmente, en este caso, se hace referencia a los procesos de negocio. Un proceso de negocio es un sistema de actividades que transforma la información y los materiales en un producto. Las cadenas de valor y los procesos de negocio producen resultados que son valorados por los clientes. Las actividades de un proceso se relacionan con otras mediante dependencias de recursos y de control.

El modelado de los procesos de negocio permite organizar la información sobre el sistema, describir cada proceso, para en una etapa posterior poder utilizar el conocimiento adquirido, aplicarlo en mejoras y servir de apoyo en la toma de decisiones.

Las técnicas más utilizadas en cuanto al modelamiento de procesos de negocios son:

Diagrama de flujo: Datan de los años 60, son una representación gráfica de la secuencia lógica de los procesos de trabajo. Puede representar operaciones, datos, direcciones de flujo y recursos.

Diagramas de flujo de datos: Son representaciones de información que permiten ver el flujo de datos a través de la organización, los procesos y transformaciones de los datos y las salidas. Sin embargo, no modela flujos de materiales, recursos, personas y otros.

Diagrama entidad-relación: Este diagrama es un modelo de red que muestra la distribución de datos almacenados en un sistema. Su foco se encuentra en los datos y sus interrelaciones.

Diagrama estado transición: Diagramas utilizados para describir la parte más dinámica de los sistemas que dependen del tiempo. Consiste en círculos que representan los estados y flechas que asemejan las transiciones entre los estados.

IDEF: Por sus siglas en inglés, *Integrated Definition por Function Modeling*, es una técnica de modelamiento que representa un tanto un modelo de procesos como una estructura de datos. Particularmente, la más utilizada para el modelamiento de procesos de negocio es IDEF0. Permite identificar los procesos y sus interfaces, así como la elaboración de documentos que permitan el control en cualquiera de las etapas de desarrollo.

Diagrama de actividad de roles: O RAD por sus siglas en inglés '*Role Activity Diagram*', se utilizan para esquematizar las actividades dentro de la responsabilidad de cada rol, la interacción entre ellos, con sucesos externos. Un rol se entiende como el comportamiento deseado de un individuo dentro de la organización.

Diagrama de interacción de roles: O RID por sus siglas en inglés '*Role Interaction Diagram*', son gráficos que muestran los roles de los procesos de negocio. Las actividades se conectan a los roles en una matriz. Se utilizan principalmente para flujos de trabajo y procesos que implican la coordinación de actividades relacionadas entre sí.

Técnica Orientada a Objetos: Se utiliza esta técnica para cuando es necesario modelar y programar procesos caracterizados como objetos, desarrollados y transformados por actividades. La metodología más utilizada es UML (Unified Modelling Language), que consiste en una visualización que permite especificar y documentar cada parte del desarrollo de software.⁹

4.3.2 Modelos de bases de datos

Se denomina a las *bases datos* como un conjunto de información agrupada que comparte un vínculo entre si y que permite su orden y/o clasificación.⁹

⁹ SANCHIS, Raquel; POLER, Raúl y ORTIZ, Ángel. Técnicas para el Modelado de Procesos de Negocio en Cadenas de Suministro. *Inf. tecnol.* [online]. 2009, vol.20, n.2 pp. 29-40.

Un modelo de base de datos es una teoría que describe cómo una base de datos es estructurada y usada. Algunos de los modelos más tradicionales son:

- Modelo jerárquico
- Modelo de red
- Modelo relacional

En un **modelo jerárquico**, los datos son organizados en una estructura parecida a un árbol, implicando un eslabón solo ascendente en cada registro para anidar, y un campo de clase para guardar los registros en un orden particular en cada lista del mismo nivel

El **modelo de red** generaliza una relación entre un conjunto de elementos y estos a su vez están compuestos de registros (que pueden ser organizados jerárquicamente). Los conjuntos definen de uno a varias relaciones entre registros.

El **modelo relacional** fue presentado por E. F. Codd en 1970 como una forma de hacer sistemas de gestión de datos más independientes.

Actualmente, los productos que utilizan bases de datos relacionales ponen en práctica un modelo que solo es una aproximación al modelo matemático definido por Codd.

En el modelo relacional se utilizan tres términos clave:

- Relaciones: Se puede ejemplificar como una tabla de con filas y columnas que permite encontrar el nexo entre un atributo y otro.
- Atributos: Es un descriptos de la relación, es el encabezado de cada una de las columnas de la tabla.
- Dominios: Conjunto de valores que puede tomar el atributo.

4.3.3 Procesos ETL

ETL proviene del inglés *Extract, Transform and Load* y es un proceso que organiza el flujo de datos entre diferentes sistemas de la organización, aportando métodos y herramientas necesarias para mover datos desde múltiples fuentes, reformatearlos y limpiarlos, cargarlos en otra base de datos, *data mart*, *data warehouse* o bien en otro sistema operacional. Esto con el fin de analizar o bien de apoyar un proceso de negocio.

10

¹⁰ http://etl-tools.info/es/bi/proceso_etl.htm

Los procesos ETL son relevantes para estudio debido a que los datos relevantes de la operación comercial se encuentran distribuidos en numerosos sistemas y funcionan bajo diferentes paradigmas, por lo que se hace necesario un trabajo de la información para lograr los resultados esperados.

Un proceso ETL consta de tres partes: Extracción, Transformación y Carga de datos.

Extracción: Consiste en extraer los datos desde los sistemas de origen. Si la extracción va a ser recurrente, ésta debe tener el menor impacto posible en los sistemas que accede, tanto a nivel volumétrico como de oportunidad de la recolección de datos. Adicionalmente, la extracción debe satisfacer en fases posteriores, las métricas y atributos que se abordan en el estudio, por lo que se debe verificar que la extracción de datos entregue lo que se requiere.

Transformación: En todo proceso ETL, esta fase es la más relevante. Esto pues acá es donde se aplican todas las lógicas de negocio y donde se crean las métricas y atributos utilizados.

Carga de datos: Para finalizar un proceso ETL, en la carga de datos se almacena la información procesada. Frecuentemente esta fase apunta a la disponibilización de la información en un reporte.

5. DESARROLLO DEL TRABAJO

5.1 Selección de principales métricas y atributos del negocio

En el marco conceptual se establecieron algunos de los conceptos de negocio, métricas y atributos utilizados en la gestión del negocio aéreo. También se dio un repaso por los modelos de pronóstico de demanda y modelos de bases de datos.

Para la construcción de un sistema que permita contener los principales indicadores es necesario establecer cuáles de ellos hoy son recurrentemente calculados por los analistas y cuáles podrían ser potencialmente agregados, sin caer en una complejidad mayor o que finalmente la capa de datos sea tal que no permita la correcta revisión.

Para levantar los indicadores se hicieron entrevistas a los diferentes analistas de negocios de RM y se fijaron los conceptos que inicialmente tendrá la visualización. Una de las consideraciones más relevantes es que independiente del diseño, la capa de datos tiene que ser dinámica y permitir posteriormente incorporar nuevas métricas y atributos si estos representan valor para el negocio:

Del total de conceptos de negocio que se revisaron, los más frecuentemente utilizados y construidos individualmente son:

- Pasajeros Reserva (*Pax Bookings*)
- Pasajeros Industria (*Pax Industry*)
 - Apertura por Carrier u operador
- Pasajeros Venta (*Pax ticket*)
- Ingresos (*Suma de Tarifas*)
- Tarifa Media

Las aperturas o atributos de estas métricas le dan un sentido de negocio, aduciendo a la utilización que tendrán, dependiendo de la profundidad de los diferentes análisis y el proceso de negocio que llevan los analistas.

Los principales elementos al momento de recopilar fueron:

- Cabina
- POS
- OD

- Tipo de pasajero
 - Identificar reservas individuales
 - Identificar reservas de grupos
- Canal de venta
 - Directo
 - Indirecto
- Mes de vuelo
- Ruta
- Anticipación de venta (En el caso de tickets)
- Anticipación de reserva (En el caso de una reserva)
- Agrupación ODPOS

La mayoría de estos elementos corresponden a definiciones conceptuales que fueron expuestas en el marco conceptual. Adicionalmente, aparecen agregaciones de fechas de vuelo (mes de vuelo), ruta y conceptos de anticipación.

Ruta es un concepto de agrupación de una tupla de vuelo, tramo y fecha de vuelo. Representa una unidad de gestión dentro de LATAM. De esta manera la agrupación de todos los vuelos que van desde **SCL** (Santiago) a **MIA** (Miami) se denomina ruta **MIA**. Esta unidad de agregación es la que finalmente agrupa ingresos y costos para ser presentada en los estados de resultados.

La **anticipación de venta** corresponde a la diferencia en días desde la fecha de emisión de un ticket hasta la fecha de salida del primer vuelo del ticket.

La **anticipación de reserva** es la diferencia en días desde la fecha de reserva y la fecha del primer vuelo de la reserva.

El **mes de vuelo**, corresponde a la agregación o jerarquía de todas las fechas de vuelo que pertenecen al mismo mes.

La **agrupación OD POS** corresponde a la agregación conjunta de diferentes orígenes, destinos y POS.

5.1.1 Análisis de métricas y atributos

Los atributos y métricas descritos previamente se llevaron a una *Matriz de Kimball* para una mejor representación:

Métricas/Atributos	Ruta	OD	Cabina	POS	Tipo de pax	Canal de venta	Mes de vuelo
Pasajeros reserva (<i>Pax Bookings</i>)	x	x	x	x	x	x	x
Pasajeros industria (<i>Pax Industry</i>)		x	x	x			x
Pasajeros venta (<i>Pax ticket</i>)	x	x	x	x	x	x	x
Ingresos (<i>Suma de Tarifas</i>)	x	x	x	x	x	x	x
Tarifa Media	x	x	x	x	x	x	x

Figura 9: Matriz de indicadores principales

En el cuadro previo, la 'x' denota la apertura del indicador a ese nivel, en caso que el indicador no pueda ser abierto por restricciones lógicas o no sea necesaria su apertura la celda aparece vacía.

5.1.1.1 Métrica Pasajero Reserva

Para la métrica **Pasajeros Reserva** fue necesario ir a la fuente *GPNR*, donde se encuentra la información de reservas de LATAM. La contabilización de los pasajeros no es algo trivial dado que la métrica abierta debe estar por **OD** y en el sistema de información por defecto la contabilización es por segmento. Esto se resuelve considerando a cada pasajero de forma única por cada OD. Es decir, si el pasajero realiza un viaje de 2 segmentos y por consiguiente la cantidad de pasajeros es 1 por cada segmento, dado que es un mismo pasajero el que realiza el vuelo la cantidad total de pasajeros por OD es 1.

Similarmente, para la contabilización de pasajeros por **cabina**, es necesario determinar la cabina del OD. Considerando que es posible viajar 2 segmentos en cabinas diferentes para una misma reserva, ¿Qué cabina tiene el pasajero en el OD?

Para unificar esta visión, se decidió que lo más representativo era utilizar la cabina del segmento que tuviera una distancia total mayor recorrida.

En el caso del atributo **ruta** este no es parte del sistema de reservas por lo que fue necesario capturarlo desde el modelo de rutas LATAM, donde se alberga esta definición.

Para el **POS**, previamente definido como el país donde se realiza la transacción, en el caso de la fuente utilizada existían problemas al extraerlo directamente, debido a que el POS correspondía al de sistema y no necesariamente reflejaba la realidad comercial. Dado esto, se consensuó con los equipos utilizar como *proxy* el código de agencia IATA¹¹. Para esto fue necesario capturar la relación entre el código de la agencia y su país desde el sistema que asigna las estructuras comerciales a la agencia (ADMC¹²). De manera similar y con la misma fuente también se identificó el **canal de venta**.

Para el **tipo de pasajero**, éstos se clasifican en pasajeros de reserva individual o de grupo. Dentro de la fuente se encuentra el *flag* que indica si el pasajero es de un tipo u otro.

Finalmente, los elementos se agrupan a nivel de **mes de vuelo**, lo que únicamente corresponde a agrupar los datos (sumatoria sobre pasajero).

La **agrupación OD POS** es una jerarquía para el origen destino por lo que se requiere de un diccionario geográfico, en LATAM denominado *Modelo geográfico*, que engloba las definiciones estándar de jerarquía desde aeropuerto a continente.

El proceso para la construcción de esta métrica queda reflejado en la siguiente figura:

¹¹ Código de agencia IATA: Código numérico único de 8 dígitos que identifica únicamente al emisor del ticket o de la reserva.

¹² ADMC: Administrador comercial

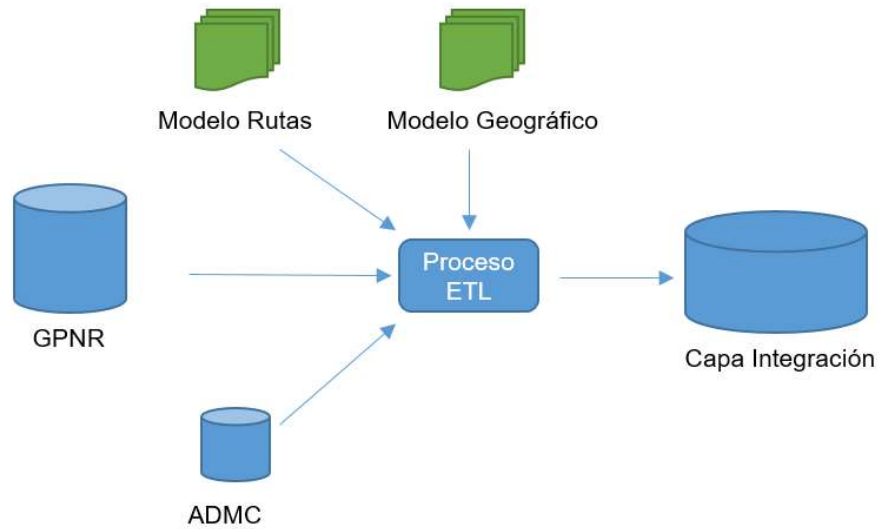


Figura 10: Modelo métrica pasajero reserva

5.1.1.2 Métrica Pasajero industria

La métrica **pasajero industria** se extrajo de la fuente MIDT. A diferencia del caso de reservas, solo es necesario tener esta información a nivel de OD, Cabina, POS y mes de vuelo. Adicionalmente, dado que los elementos a comparar luego serán por línea área, se debe contar también con ese atributo.

Dado que no es necesario generar el indicador a nivel de ruta la construcción es más directa quedando como:

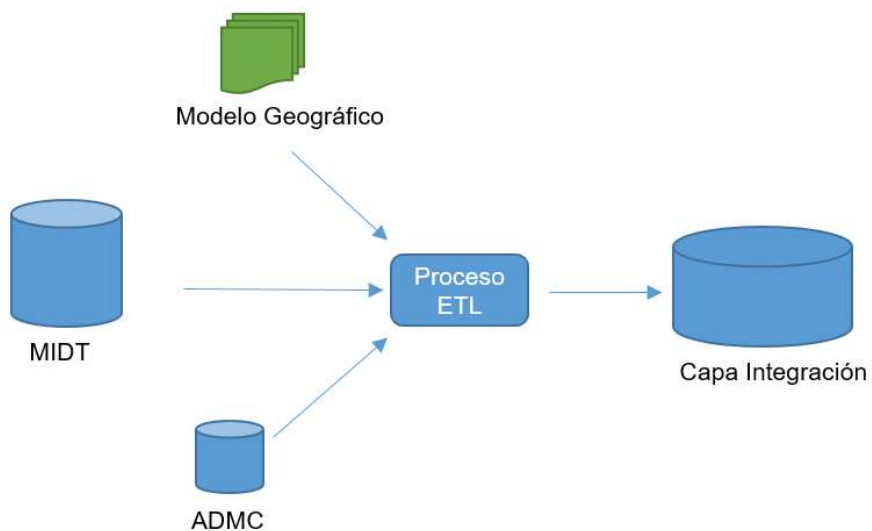


Figura 11: Modelo métrica pasajero industria

5.1.1.3 Métrica Pasajero venta, ingreso y tarifa media

Las métricas de **pasajero venta**, **ingreso** y **tarifa media**, se obtienen desde la fuente JURO/TCN. En esta base se encuentran los tickets. Estos elementos tienen una lógica estándar en la industria, cada elemento ticket puede tener hasta cuatro cupones que representan los diferentes segmentos de vuelo. Esta es una restricción histórica que se refleja también en los sistemas y cuando un mismo pasajero desea hacer más de 4 segmentos en un mismo viaje se debe generar un ticket adicional denominado *ticket en conjunción*. Para evitar esta problemática, el sistema se diseñó con un código denominado *master ticket* que almacena todos los tickets y sus cupones, eliminando la limitante. De esta manera:

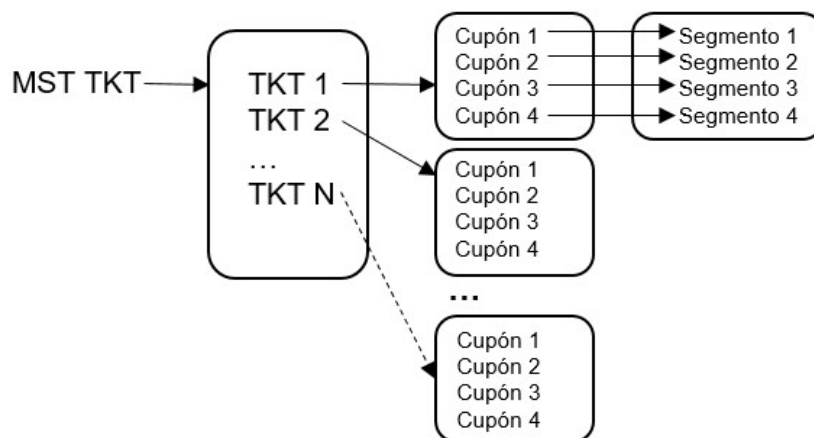


Figura 12: Relación ticket y cupón

Para **pasajero venta**, siguiendo la definición de OD estándar, se asignará de la misma forma que en la métrica pasajero reserva, contabilizando una vez a cada pasajero por OD. Así se mantiene la integridad y coherencia en la información.

Para el **ingreso** se realiza la suma sobre todos los valores tarifarios por cupones/segmento por OD (se agrupan por segmentos pertenecientes a mismo OD). Para la ruta el tratamiento es diferente dado que, a diferencia de los pasajeros, la interpretación de negocio del ingreso puede ser fraccionada a rutas diferentes, es decir no tiene sentido hablar de 0,5 pasajeros por ruta, pero si un mismo pasajero realiza un

OD utilizando 2 rutas, el monto total si puede ser distribuido y asignado a cada una, este elemento de repartición se denomina prorrato y es común dentro del negocio aéreo.

En este caso, el prorrato de los ingresos para los distintos niveles será realizado de acuerdo a las definiciones estándar de la industria. Para esto se utiliza la distancia SRP¹³. Esta 'Distancia' permite distribuir los ingresos de manera tal de relativizar las grandes diferencias entre distancias geográficas para 2 o más segmentos.

La **tarifa media** corresponde al ingreso prorrato o agrupado (según sea ruta o OD) dividido por el número de pasajeros.

Finalmente, el modelo para estas últimas métricas queda de la siguiente forma:

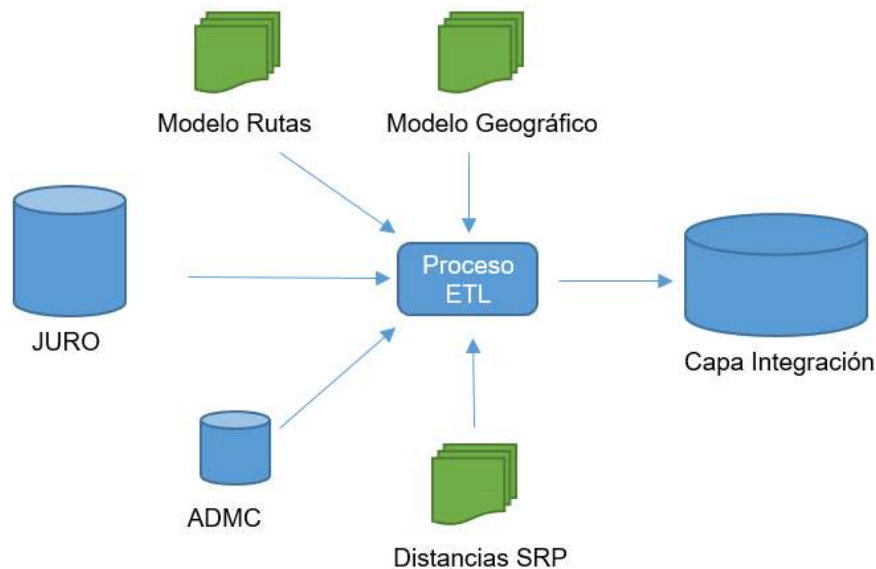


Figura 13: Modelo métrica pasajero venta, ingreso y tarifa media

5.1.1.4 Consideraciones generales para métricas y atributos

Las siguientes consideraciones son globales para todos los atributos mostrados previamente:

¹³ SRP: Straight Rate Proration

- Los tickets son elementos posteriores dentro del ciclo de venta por lo que cuando se compara los pasajeros venta y reserva siempre se cumple la relación:

$$\sum Pax\ Venta \leq \sum Pax\ Reserva$$

- Se debe filtrar previamente a la capa de integración los pasajeros que no se gestionan directamente por los equipos comerciales en estudio, estos pasajeros corresponden a los liberados (viajes por beneficio empresa) y pasajeros Award (referentes al programa de pasajero frecuente de la compañía)
- Para integrar una misma visión comercial de la información debe existir coherencia en relación a la 'foto' de la realidad que se está exponiendo, este concepto se denomina *Fecha vista*, y corresponde al día de observación. De esta manera se alinean las distintas fuentes para tener una forma clara y transparente de determinar la fecha a la cual corresponde la observación que se está trabajando.

5.2 Proceso ETL

5.2.1 Extracción de datos

Desde cada una de las bases y tablas paramétricas que se revisaron en el capítulo anterior se construyeron las ETL de extracción, lógica de negocio y carga de datos. Esto se realizó en una herramienta open-source denominada *Pentaho Data Integration* (PDI, Detalles en Anexos), la cual se eligió por sobre otras soluciones por 3 razones principales:

- Herramienta ya es utilizada en la empresa, por lo tanto, facilita el traspaso futuro del proyecto y los empleados ya están familiarizados con ésta.
- Dado que es una herramienta *open source*, el costo de implementación se reduce considerablemente versus otras opciones.
- PDI permite múltiples tipos de input, incorpora los drivers para todos los tipos de fuentes que se requieren, lo que facilita la conexión con los servidores donde se almacena la información.

Según las reglas descritas previamente, se programaron las extracciones de tal forma que las transacciones se realizaran automáticamente y autónomamente, de esta manera, esto puede operar de sin requerir de personas para ejecutar ni recolectar la información.

5.2.2 Transformación de datos y carga

La parte siguiente a la obtención de datos, corresponde a la transformación, aplicando las reglas de negocio y consideraciones previamente descritas. Esto se aplica sobre la Capa de integración, donde se construyeron los procesos que agregan la data al nivel que será usado luego por los analistas. En esta parte están consideradas las lógicas de negocio descritas en el capítulo previo.

También es necesario disponer de los datos para la generación de la proyección, lo que será considerado como uno de los outputs de la capa de integración, pero debe ir sobre las tablas intermedias temporales ya que es necesario un nivel de desagregación mayor. De forma gráfica esto lo podemos visualizar como:

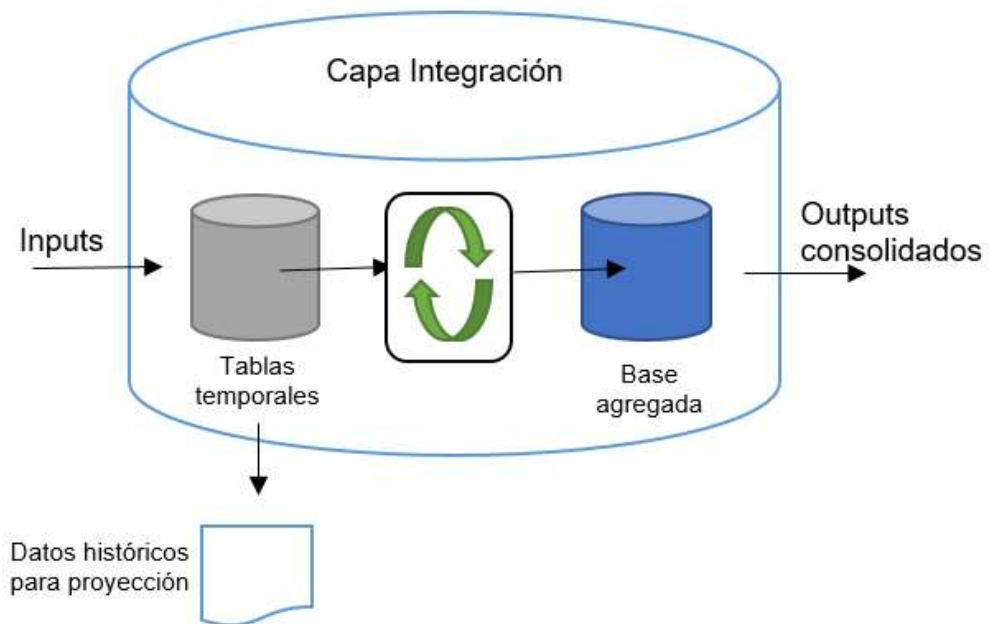


Figura 14: Capa integración

Tanto las tablas temporales como la base agregada se encuentran en un servidor de base datos MySQL, el cual se encontraba disponible para la realización de este trabajo.

5.3 Requerimientos de negocio asociados a la proyección

5.3.1 Información y procedimiento

La proyección de demanda de pasajeros está asociada a las reservas tanto de canal directo como indirecto. Para determinar la demanda futura de tickets, se cuenta con información histórica desde el año 2007 de venta de tickets y de las reservas de éstos. Dado que los tickets son una realización posterior a la reserva se utilizarán exclusivamente las reservas de pasajeros, dado que los tickets obedecen a otras condiciones. Para esta parte del trabajo se denominará reserva o ticket de manera indistinta.

El detalle de la proyección se requiere desagregado a nivel de OD (Origen-Destino pasajero), POS y Cabina. La frecuencia con se requiere es semanal, que corresponde a la actualización con que se lleva a cabo el proceso comercial de RM. Por lo tanto, cada semana se irán proyectando los meses vuelo futuro.

Por ejemplo: Para el ODPOS: SCLMIACL Cabina Y (Origen: Santiago, Destino: Miami, POS: Chile, Cabina Y o económica) la demanda de tickets hasta la primera semana de enero fue 523 reservas acumuladas para volar dentro del mes de marzo, la segunda semana estas reservas aumentaron a 624, luego la tercera semana de enero se tienen 754, etc. De esta forma se tiene una evolución semana a semana de la demanda futura para volar dentro de marzo.

Dado lo anterior la salida de información requerida para el análisis contempla las siguientes aperturas:

- OD
- POS
- Cabina
- Fecha proceso (fecha vista semanal)
- Mes de vuelo

El horizonte de proyección es de los siguientes 6 meses, esta decisión se basa en que los meses de vuelo posteriores a este periodo no tienen aún una gestión estructural semana a semana.

Adicionalmente, se requiere de la apertura del canal de venta por los conceptos agrupados

Pax reservas canal Directo	Pax reserva individual
	Pax reserva grupo
Pax reservas Canal Indirecto	LAN
	Otras aerolíneas

Figura 15: Canal de reserva

Este detalle de proyecciones es coherente con la gestión de negocio ya que algunas de estas aperturas son revisadas por otros equipos de gestión diferentes a RM, para detectar nuevas oportunidades o cuando se visualizan caídas y levantar problemas en caso que sea necesario

5.3.2 Performance herramienta visualización

La capacidad de poder responder con rapidez a las consultas para los distintos filtros que se escojan es un desafío importante, dado que el sistema debe albergar gran cantidad de información.

Se establece que el tiempo de respuesta debe ser como máximo 1 minuto, para cada una de las vistas de negocio, esto para poder gestionar de manera adecuada la información obtenida.

5.3.3 Proceso de proyecciones

El proceso de proyecciones es realizado por los analistas de RM y es parte del ciclo donde se agrega valor de forma táctica. Los analistas van semana a semana pronosticando pasajeros con la siguiente finalidad:

- *Mejorar ajuste de parámetros de PROS:* Esta herramienta es donde se integran y definen las clases tarifarias y los cupos asociados a los vuelos, teniendo una visión adicional de la demanda futura es posible configurar de mejor manera la oferta, al poder segmentar con mayor precisión la demanda. Si no se manipula correctamente PROS, los vuelos pueden cerrar antes de tiempo o bien puede venderse de forma subóptima. Si el vuelo cierra antes de tiempo hay un costo por no haber vendido tickets de mayor valor o por el contrario queden asientos vacíos en el avión.
- *Gestión mirando a futuro de otras áreas:* Con las proyecciones de demanda a futuro se visualizan oportunidades con otros equipos para el entendimiento del mercado, en especial para fluctuaciones importantes que alerten sobre posibles problemas u oportunidades.

- *Definición del presupuesto:* Las proyecciones son utilizadas para definir el presupuesto anual, que se va actualizando trimestralmente.

5.3.4 Proyección de demanda y reservas históricas

Uno de los elementos críticos del sistema es la confiabilidad de las proyecciones de demanda, considerando las reservas históricas por volar como base para predecir la demanda futura.

Las reservas realizadas en el pasado se denominan históricas. Esto quiere decir que fueron adquiridas previamente y puede que hayan sido voladas o serán voladas en fechas futuras, dependiendo de su anticipación. Para el caso de las reservas por volar, éstas caen dentro del periodo de tiempo cuya demanda interesa proyectar. Entonces, en la predicción, es posible conocer parte de la demanda. A diferencia de un producto de consumo inmediato, donde la compra y la ejecución del servicio se dan al mismo tiempo, para el caso del transporte aéreo el modelo implica una diferencia entre ambos eventos.

Por ello, se estudiaron diversos modelos, principalmente los que consideraran el conocimiento de las reservas históricas por volar, que proveen de más información al momento de hacer el pronóstico.

El hecho de contar con la información de reservas históricas permite conocer parte de la demanda previamente, y su valor final se va revelando a medida que avanza el tiempo.

Las reservas se realizan para una fecha y hora determinada con origen y destino específico, conteniendo también los detalles de la cantidad de pasajeros, POS, el **routing**, entre otros. Todo este detalle es capturado en el sistema en la capa de integración presentada previamente.

Como parte de los requerimientos, los pronósticos se realizarán por ODPOS, lo cual se denomina comúnmente como *mercado* en la compañía.

Para poder construir los mercados y las líneas de tiempo históricas, se suman todos los pasajeros de las reservas que tengan el mismo origen y destino y el mismo POS. Respecto a la periodicidad o fechas vistas, se agrupan las ventas de forma semanal, esto quiere decir que cada semana se “saca una foto” del estado de las reservas. A continuación se muestra una tabla explicativa de este procedimiento:

Origen - Destino	SCL - BUE
Fecha	02-sep-09
Vuelo	LA785
Hora	17:00
Semana / Tickets	
1	5
2	15
3	23
4	48
5	56
6	75
7	82
8	98
9	103
10	117
11	123
12	167
13	179
14	195

Figura 16: Cuadro comparativo semanas

La tabla muestra el detalle de un vuelo histórico considerando 14 semanas previas, en donde la semana 1 es la más alejada a la fecha de vuelo y por lo tanto es la que tiene menor cantidad de tickets o reservas acumuladas.

El vuelo con origen Santiago (SCL) y destino Buenos Aires (BUE), para una fecha de vuelo determinada (02-septiembre-2009) corresponde al vuelo LA785 con partida a las 17:00hrs. A medida que avanzan las semanas se van acumulando las reservas hasta llegar a la semana 14, que corresponde a la semana del vuelo. Por lo tanto, esta acumulación de reservas describe una curva creciente:

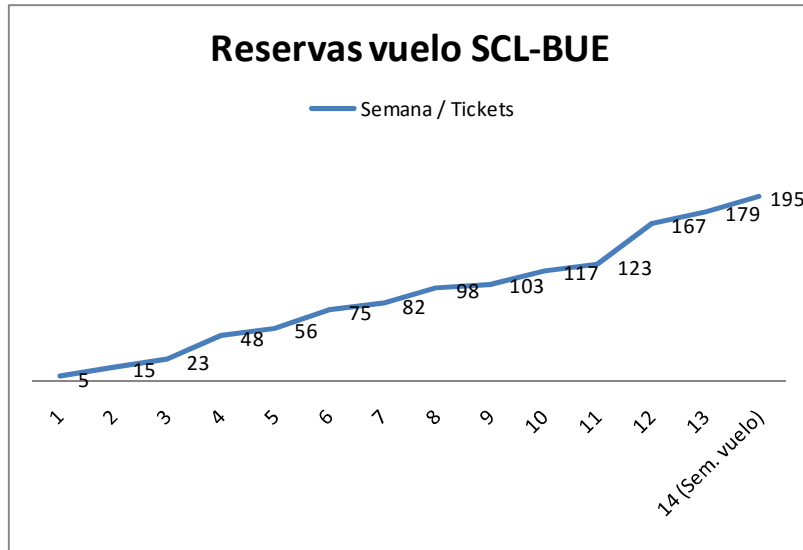


Figura 17: Grafico reservas mercado SCL-BUE

El grafico muestra las reservas acumuladas para un vuelo determinado, sin embargo, no se hace diferencia del *Point of Sale* (POS) o en qué país se vendieron los tickets. Para esto se identifica en un vuelo los tickets correspondientes a un POS determinado. En el ejemplo anterior, los tickets vendidos en Chile (CL) representan el 80% de la venta. Gráficamente:

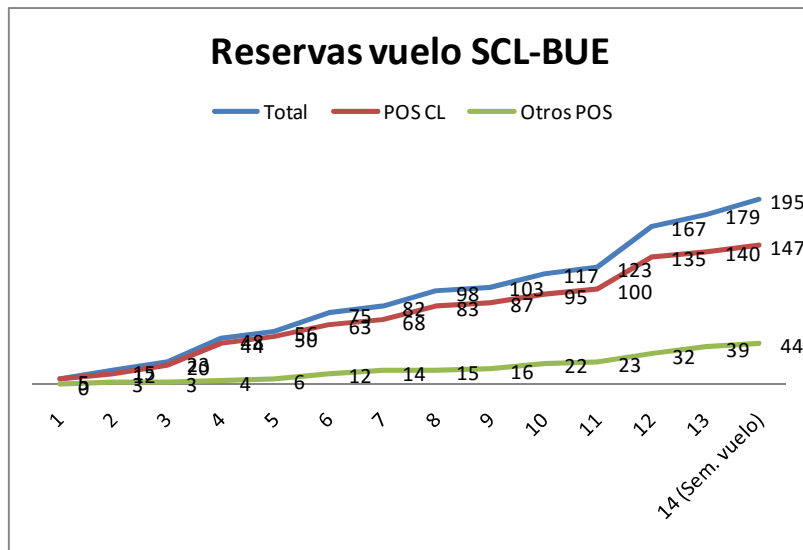


Figura 18: Grafico reservas mercado SCL-BUE por POS

En este grafico se observa que las reservas originadas en Chile representan el mayor porcentaje dentro de las reservas del vuelo, mientras que la curva "Otros POS" acumula las reservas originadas en otros países. El mayor porcentaje de reservas vendidas en

Chile dentro del vuelo se debe a que el origen del viaje es en Santiago, por lo tanto, lo más probable es que la mayor cantidad de boletos hayan sido comprados en el país de origen del vuelo.

Finalmente, como input del modelo de proyecciones, se agrupan las reservas de los vuelos con el mismo Origen y Destino (con fecha de salida dentro del mismo mes) y que tengan el mismo punto de venta (POS).

Llevado al ejemplo anterior, se suman todas las reservas de todos los vuelos de SCL-BUE con fecha de salida en “Septiembre 2009” que sean vendidos en Chile. Luego se agrupan todos los vuelos cuyo mes de vuelo sea Septiembre.

Considerando que LATAM ofrece 300 destinos posibles, por lo tanto se estima que la red completa puede contener hasta 90.000 OD. Al abrir los OD por POS entonces es posible considerar hasta 4.500.000 mercados a proyectar en el estudio (90.000 por 50 POS)

5.3.5 Pickups Multiplicativos en las reservas

Los *pickups* multiplicativos son la forma en que se agrupan las reservas para describir las tasas de cambio entre diferentes semanas de la curva de reservas acumuladas. Siguiendo con el caso de ejemplo, se tiene lo siguiente:

Origen - Destino	SCL - BUE	
Fecha	02-sep-09	
Vuelo	LA785	
Hora	17:00	
Semana / Tickets		Pick up
1	5	1
2	15	3
3	23	1,53333333
4	48	2,08695652
5	56	1,16666667
6	75	1,33928571
7	82	1,09333333
8	98	1,19512195
9	103	1,05102041
10	117	1,13592233
11	123	1,05128205
12	167	1,35772358
13	179	1,07185629
14	195	1,08938547

Figura 19: Cuadro de Pick-Ups

La columna *Pick up* representa la tasa de cambio entre cada una de las semanas de evoluciones de las reservas acumuladas. Por ejemplo, entre la semana 1 y 2 hubo un aumento de 3 veces la magnitud de las reservas, entre la semana 2 y 3 hubo un aumento de 1,5 veces las reservas, etc.

De este modo, se tiene un avance en proporción, no importando la magnitud total del cambio en la cantidad de reservas. Luego, agregando los datos de todos vuelos de un mes y filtrando por POS, se tienen curvas de pickups por cada uno de los 4.500.000 ODPOS en estudio. Para el caso de las reservas de la industria, este número es aún mayor, dado que ahí se encuentran todos los mercados globales

5.3.6 Matrices de anticipación

Una vez obtenidos los pickups y las reservas se pueden construir las matrices de anticipación, a las que finalmente aplicaran los modelos de proyecciones. Estas matrices se representan de la siguiente manera:

Mes Vuelo / Pickups	Semanas anticipacion							
	1	2	3	4	5	6	7	8
ago-09	0,64	0,29	1,02	1,89	2,08	2,16	2,71	1
oct-09	0,32	0,29	1,01	2,62	1,35	3,58	6,76	1
nov-09	0,34	0,19	1,27	2,92	4,60	7,02	11,13	1
dic-09	0,19	0,22	1,43	1,68	2,62	4,96	7,67	1
ene-10	0,16	0,17	1,07	3,09	1,50	3,24	4,30	1
feb-10	0,98	1,86	1,91	2,65	2,49	2,95	6,68	1
mar-10	0,31	0,54	1,99	1,83	3,46	7,47	14,36	1
abr-10	0,03	0,05	1,04	3,01	6,18	2,38	4,89	1
jun-10	0,45	0,39	1,12	1,41	1,21	2,21	2,20	1
jul-10	0,77	0,88	1,60	3,71	3,26	5,64	9,37	1
ago-10	0,55	0,16	1,01	2,25	4,73	1,95	4,89	1
sep-10	-	-	-	-	1,55	2,75	1,51	1
oct-10	-	-	-	-	-	-	-	1

Figura 20: Cuadro por anticipación y pickups

Cada uno de los mercados estudiados tiene una matriz de anticipación que contiene los pickups para cada mes cerrado y su evolución desde el pasado.

Por ejemplo, en la tabla para ene-10 (enero 2010), esta descrito su comportamiento por estos factores de cambio con 8 semanas de anticipación.

En las proyecciones realizadas se utilizaron para cada mes de vuelo, 25 semanas de anticipación, lo que implica que cada uno de los ODPOS tiene 25 series de tiempo.

Cada uno de los valores de la matriz es el aumento en tasa de una semana a otra.

5.3.7 Modelos de proyecciones

Se decidió utilizar las series de tiempo en conjunto con los pick-ups multiplicativos debido a que el modelo matemático resultante ofrece una explicación del comportamiento del mercado. Sin embargo, dada la cantidad de series de tiempo, el estudio particular para encontrar el mejor ajuste puede ser complejo.

Debido a que todos los mercados presentan diferencias en el comportamiento, pretender resolver el problema con un solo modelo no es factible. Dada esta situación, se decidió aplicar todos los modelos de series de tiempo al histograma de picks ups y luego, en función del error de cada uno, buscar el que mejor se ajuste a la curva.

De este modo, una vez que todos se aplican a cada serie de tiempo, se auto-elige el modelo que mejor se ajusta a la curva histórica, siguiendo como métrica para decidir entre un modelo y otro la desviación absoluta de la media (MAD). Este funcionamiento de los modelos de proyección puede esquematizarse de la siguiente manera:

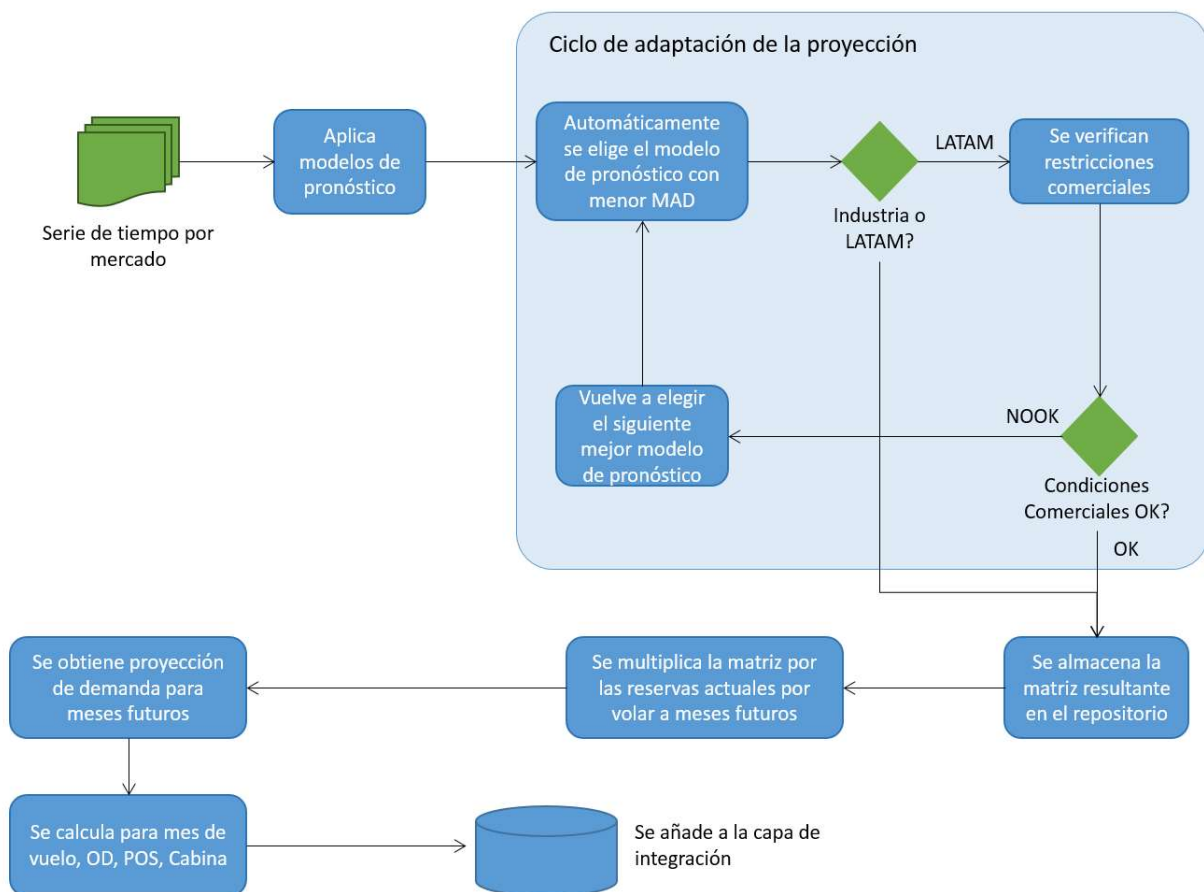


Figura 21: Proceso cálculo proyección

El input que recibe el modelo de proyecciones es el de las matrices de anticipación, estas contienen los saltos semana a semana para las reservas históricas.

Luego se aplican todos los modelos de pronóstico revisados en el marco conceptual. Para todos ellos se calcula la desviación y el modelo que menos se desvía es el elegido.

Posteriormente, para la información de LATAM, se verifican las restricciones comerciales. Es decir, que la proyección no supere la capacidad u oferta, esto es así dado que se está proyectando la demanda *constrained*.

Para el caso de pasajeros industria esto no aplica dado que no se tiene información de la capacidad u oferta de la industria para los mercados.

El output del algoritmo de proyecciones son las matrices de pickups para el futuro o, en definitiva, la proyección en los cambios de reservas a futuro. Esto corresponde al complemento de la matriz de anticipación de la figura 19:

Mes Vuelo / P	Semanas anticipacion							
	1	2	3	4	5	6	7	8
ago-10	-	-	-	-	-	-	-	1
sep-10	0,34	0,76	0,45	0,23	-	-	-	1
oct-10	0,56	0,67	0,49	0,32	0,78	0,56	0,12	1

Figura 22: Matriz resultado pickups

Finalmente, esta matriz se multiplica por la cantidad de pasajeros reserva que ya se encuentran en el sistema para meses de vuelo futuros. De esta manera tenemos el número de pasajeros reserva proyectados.

Para el desarrollo de esta herramienta se utilizó Java y una librería de código abierto para la aplicación de series de tiempo llamada *open forecast*. La característica principal de la metodología implementada es que no requiere de un estudio particular caso a caso de cada mercado, tampoco de la intervención de personas por lo que los resultados pueden ser evaluados directamente cuando las métricas estén construidas en el reporte.

5.4 Solución como sistema de información

5.4.1 Elección de herramienta visualización y funcionamiento

Un desafío importante en la construcción de sistemas de información es la capa de cara al usuario. El mercado ofrece muchas alternativas para que una vez que los datos estén cargados, los usuarios puedan explorar, construir elementos visuales, gráficos y reportes.

La visualización de la herramienta es crítica porque finalmente por más valor que tenga la información, si el “look and feel” del reporte no es el adecuado puede desacreditar el trabajo que hay atrás.

Dado las condiciones anteriores, se decidió, para este proyecto, explorar una solución que permitiera conservar Excel como principal herramienta para el trabajo de datos, de este modo la integración y adaptación al día a día de la gestión sería más exitosa (*Change Management*). Una de las herramientas que respondió a los puntos previos es PALO, software propietario de la empresa JEDOX.

La versión *Community* utilizada en este trabajo permitía además trabajar con grandes volúmenes de información, por lo cual las restricciones de espacio que tradicionalmente tienen las herramientas de hojas de cálculo desaparecen, dado que PALO funciona con un motor OLAP en memoria, permitiendo consultar un cubo de información con gran cantidad de datos.

Para poder almacenar los cubos de información, los datos deben venir agrupados al nivel de presentación del reporte. Por ello, se diseñan las bases de datos del modelo de integración de la forma adecuada para alimentar los cubos de información.

Una representación gráfica de este relacionamiento es:

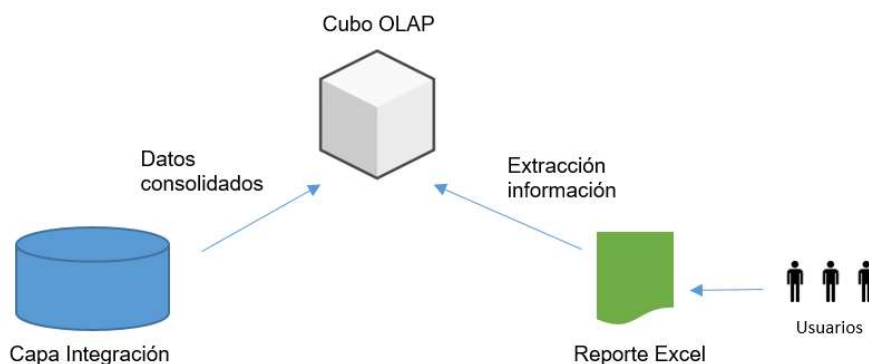


Figura 23: Modelo de visualización

En la imagen, se observa que la capa de integración, la cual contiene los datos ya consolidados al nivel de agregación requerido, carga la información en el cubo OLAP. Posteriormente Excel realiza la extracción de la información para cargar en el reporte y ser consumida por los usuarios.

5.4.2 Presentación visual de la herramienta

La presentación de la herramienta es en formato Excel, pero está conectado a un servidor que posee un cubo OLAP con la información consolidada. La apariencia del reporte es de un archivo Excel tradicional, pero al modificar los filtros se ejecuta una consulta al motor OLAP.

Visualmente, se puede separar la herramienta en 4 grandes grupos

Vista POS

The screenshot shows an Excel spreadsheet with a filter panel at the top and a data table below. The filter panel includes 'Fecha Proceso' (15-08-2010), 'Mes de Vuelo' (2010-8), and 'Cabina' (Y). The data table is titled 'Evolución Booking Totales' and shows sales data for various countries (Pos/Item) across multiple time periods (TOT_AY, TOT_LYMF, etc.).

Pos/Item	Total Ingresos	Total Proyecciones	Evolución Booking Totales						Ev	
	INGRESOS	PROY_TD	TOT_AY	TOT_LYMF	TOT_LY01M	TOT_AYLV	TOT_LYLV	AY	LYMF	
Total	248,183,815	1,479,049	1,526,566	1,305,117	1,435,512	1,379,658	1,172,565	774,396	639,247	
CL	61,974,529	454,690	481,177	404,929	452,524	425,724	353,730	346,163	279,557	
AR	43,673,107	260,762	239,217	227,047	249,539	216,907	201,598	97,524	95,194	
US	30,446,813	105,898	112,497	94,681	93,772	108,253	92,627	20,413	21,744	
PE	31,181,825	336,581	364,524	289,475	338,723	315,762	245,211	233,327	169,089	
ES	16,466,099	48,344	44,182	42,205	40,996	43,642	42,783	6,818	8,017	
BR	11,800,296	42,974	57,178	27,802	27,370	54,728	27,612	6,228	3,484	
EC	10,028,072	75,206	69,018	63,150	77,396	60,378	55,548	38,254	32,714	
VE	6,911,350	21,154	21,719	26,904	26,508	21,350	27,413	4,002	4,213	
AU	5,003,169	11,903	8,821	11,989	11,847	9,084	12,106	631	2,368	
MX	4,557,709	13,629	16,713	12,445	12,744	15,803	12,130	2,692	1,725	
JP	1,266,544	4,237	4,271	4,207	3,999	4,470	4,515	4	0	
DE	3,520,918	16,217	15,547	16,961	16,581	15,127	16,884	4,490	6,623	
IT	3,980,476	11,509	12,611	7,902	7,617	12,336	8,018	60	126	
FR	2,840,496	10,637	9,836	8,754	8,647	9,549	8,796	541	273	
CO	2,656,044	13,096	14,275	13,680	13,881	13,627	12,240	4,779	4,146	
GB	2,332,023	9,988	7,035	8,634	8,764	8,954	8,379	322	3,147	
UY	2,266,743	10,285	10,684	10,472	10,521	10,157	10,086	3,445	3,079	
CA	1,519,955	7,830	8,324	7,481	7,481	7,998	7,113	1,074	1,041	

Figura 24: Herramienta / Vista POS

Esta vista entrega una visión global de venta para todos los POS. En la parte superior se encuentran los filtros:

- Fecha proceso: Corresponde a la fecha vista (Última actualización)
- Mes de vuelo
- Cabina

El detalle de cada columna en la herramienta es el siguiente:

Pos/Ítem: Contiene los POS ordenados por mayor ingreso, independiente del origen-destino que posea, para los filtros del ejemplo; Chile (CL) representa el más alto ingreso.

“Total ingresos” - INGRESOS: Suma de los ingresos en dólares proyectados, este resultado corresponde a los pasajeros proyectados multiplicados por la tarifa media de cada POS.

“Total proyecciones” - PROY_TD: Es la proyección de la demanda de pasajeros reserva agrupada de la demanda.

Evolución Bookings Totales: Se subdivide en 5 columnas que entregan resultado dependiendo de las fechas vista y vuelo consideradas:

TOT_AY: Total de pasajeros (*bookings*) hasta la fecha de proceso especificada para el mes de vuelo correspondiente.

TOT_LYMF: Total de pasajeros reservas que se llevaban a la misma fecha del año anterior.

TOT_LYOUT: Total de pasajeros reserva para el mes de vuelo equivalente del año pasado.

TOT_AYLW: Total de pasajeros reserva la semana previa de este año para el mes de vuelo.

TOT_LYLW: Total de pasajeros reserva la semana previa a la equivalente del año pasado.

Adicionalmente a las columnas totales (evolución de bookings totales), también está presentes el detalle desagregado por cada uno de los canales con los mismos indicadores previos.

Vista Agrupación POS

Agrupación/Item	Total Ingresos	Total Proyecciones	Evolución Booking Totales						Evolución Bookings Inc		
	INGRESOS	PRDY_TD	TOT_AY	TOT_LYMF	TOT_LYOUT	TOT_AYLW	TOT_LYLW	AY	LYMF	LYOUT	
US	30,446,813	105,898	112,497	94,681	93,772	108,253	92,627	20,413	21,744	21,402	
(US) HYC-PE	2,214,121	5,924	6,117	5,728	5,648	6,103	5,686	687	1,430	1,388	
(US) MIA-PE	2,005,066	7,452	7,218	6,189	6,015	7,114	5,823	1,784	2,396	2,325	
(US) GYE-HYC	2,373,795	7,386	7,920	9,024	8,811	7,740	9,175	1,546	1,841	1,772	
(US) HYC-GYE	2,373,795	7,386	7,920	9,024	8,811	7,740	9,175	1,546	1,841	1,772	
(US) REGIONAL	1,726,734	8,040	8,784	6,003	6,158	7,866	5,674	520	417	436	
(US) CL-HYC	685,243	1,288	1,376	1,146	1,126	1,284	1,132	300	353	334	
(US) HYC-CL	685,243	1,288	1,375	1,145	1,125	1,283	1,131	299	353	334	
(US) LAX-PE	1,394,381	4,067	3,657	4,431	4,303	3,515	4,335	1,104	1,545	1,512	
(US) MIA-AR W/F	995,650	2,068	1,491	1,939	2,014	1,428	1,901	394	751	751	
(US) BUE-MIA	909,632	1,952	1,410	1,862	1,937	1,345	1,818	383	726	725	
(US) CL-MIA	617,517	1,197	1,218	1,082	1,106	1,122	1,013	315	313	306	
(US) MIA-CL	617,517	1,197	1,214	1,079	1,104	1,122	1,011	317	313	306	
(US) IIO GATEWAY-PE	961,122	4,042	3,820	3,062	3,005	3,751	3,043	312	275	265	
(US) PE-PE	675,931	5,144	5,908	4,675	4,716	5,622	4,495	568	312	312	
(US) MIA-IIO	726,627	2,497	2,534	1,993	1,941	2,471	1,943	900	889	870	
(US) AR-LAX	476,597	1,189	1,076	1,125	1,104	1,062	1,120	207	329	319	
(US) OTROS LATAM-US	618,700	1,702	1,838	1,014	1,032	1,763	971	208	155	147	
(US) GYE-MIA	489,002	1,758	1,286	1,260	1,200	1,219	1,279	453	715	675	

Figura 25: Herramienta / Vista agrupación OD POS

Esta vista muestra el detalle por grupo de OD POS.

Las agrupaciones son conjuntos de OD POS para la gestión comercial, provenientes de un trabajo inter áreas y permiten agrupar mercados con comportamientos comunes y definir focos de gestión.

Es una vista en detalle de la hoja anterior (POS) ya que en el filtro se elige en específico el POS que se quiere detallar.

Los filtros para esta vista son los siguientes:

- Fecha proceso: Corresponde a la fecha vista (Última actualización)
- Mes de vuelo
- Cabina
- POS

El detalle de cada columna en la herramienta es el siguiente:

Agrupación/ítem: Lista de Agrupación de OD POS

“Total ingresos” - INGRESOS: Suma de los ingresos en dólares proyectados, este resultado corresponde a los pasajeros proyectados multiplicados por la tarifa media de cada Agrupación OD POS.

“Total proyecciones” - PROY_TD: Proyección agrupada de la demanda

Evolución Bookings Totales: Se subdivide en 5 columnas que entregan resultado dependiendo de las fechas vista y vuelo consideradas:

TOT_AY: Total de pasajeros (*bookings*) hasta la fecha de proceso especificada para el mes de vuelo correspondiente.

TOT_LYMF: Total de pasajeros reservas que se llevaban a la misma fecha del año anterior.

TOT_LYOUT: Total de pasajeros reserva para el mes de vuelo equivalente del año pasado.

TOT_AYLW: Total de pasajeros reserva la semana previa de este año para el mes de vuelo.

TOT_LYLW: Total de pasajeros reserva la semana previa a la equivalente del año pasado.

Vista ODPOS

OD Item	Ingresos		Total Proyecciones	Evolución Booking Totales								
	INGRESOS	TARIFA_MEDIA	PROY_ID	TOT_AY	TOT_LYMF	TOT_LYOUT	TOT_AYLW	TOT_LYLW	AY	LYMF	LYOUT	A
(US) CL-LAX	252,309		555	648	579	577	653	587	246	241	240	
LAXSCL	243,584	476	512	593	507	504	598	513	229	222	222	
LAXLSC	3,067	253	12	12	10	10	13	9	1	2	2	
HOGLAX	2,468	359	7	7	8	8	6	9	2	4	4	
AHFLAX	1,390	278	5	6	4	6	6	5	1	0	1	
LAXPMC	1,332	444	3	5	5	3	4	5	3	3	1	
IPCLAX	468	468	1	3	13	13	5	13	0	0	0	
LAXZCO	0	0	1	2	1	1	2	2	2	1	1	
OTROS (US)CL-LAX	0	0	3	0	9	9	0	10	0	1	1	
BBALAX	0	0	0	0	0	0	0	0	0	0	0	
LAXPUQ	0	0	0	0	0	0	0	0	0	0	0	
CCPLAX	0	0	8	8	9	9	7	8	2	5	5	
ABILAX	0	0	0	0	0	0	0	0	0	0	0	
KIALAX	0	0	3	12	13	14	12	13	6	3	3	

Figura 26: Herramienta / Vista ODPOS

Esta vista corresponde a la visión en detalle del origen destino, los filtros en esta vista son los siguientes:

- Fecha de proceso: Corresponde a la fecha vista (Actualización más reciente)
- Mes de vuelo
- Cabina
- Agrupación de OD POS

El detalle de cada columna en la herramienta es el siguiente:

“Ingresos”

INGRESOS: Suma de los ingresos en dólares proyectados, este resultado corresponde a los pasajeros proyectados multiplicados por la tarifa media de cada OD.

TARIFA MEDIA: Valor de la tarifa media de los pasajeros tickets que ya realizaron la emisión del ticket venta.

“Total proyecciones” - PROY_TD: Proyección agrupada de la demanda

Evolución Bookings Totales: Se subdivide en 5 columnas que entregan resultado dependiendo de las fechas vista y vuelo consideradas:

TOT_AY: Total de pasajeros (*bookings*) hasta la fecha de proceso especificada para el mes de vuelo correspondiente.

TOT_LYMF: Total de pasajeros reservas que se llevaban a la misma fecha del año anterior.

TOT_LYOUT: Total de pasajeros reserva para el mes de vuelo equivalente del año pasado.

TOT_AYLW: Total de pasajeros reserva la semana previa de este año para el mes de vuelo.

TOT_LYLW: Total de pasajeros reserva la semana previa a la equivalente del año pasado.

Otras Vistas

Rutas: Vista de apoyo a la consolidación de la información. Esta sección entrega información global por routing.

Gráficos: Vista de apoyo a la consolidación de información. Otorga una mirada visual de las vistas anteriores.

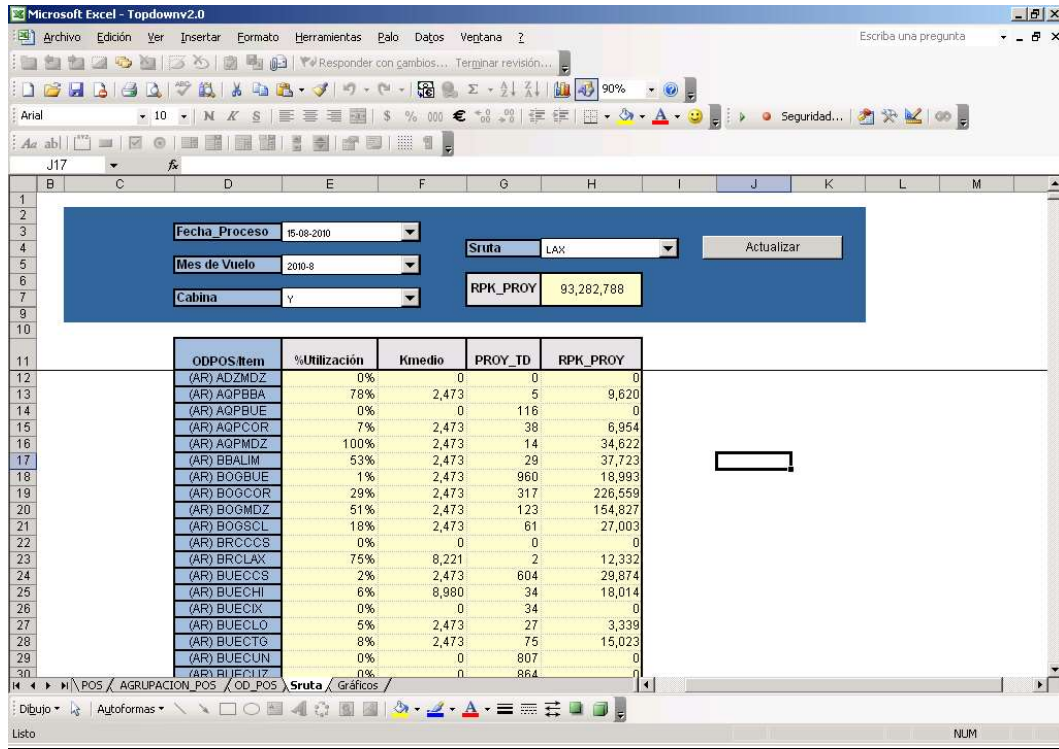


Figura 27: Herramienta / Vista Sub ruta

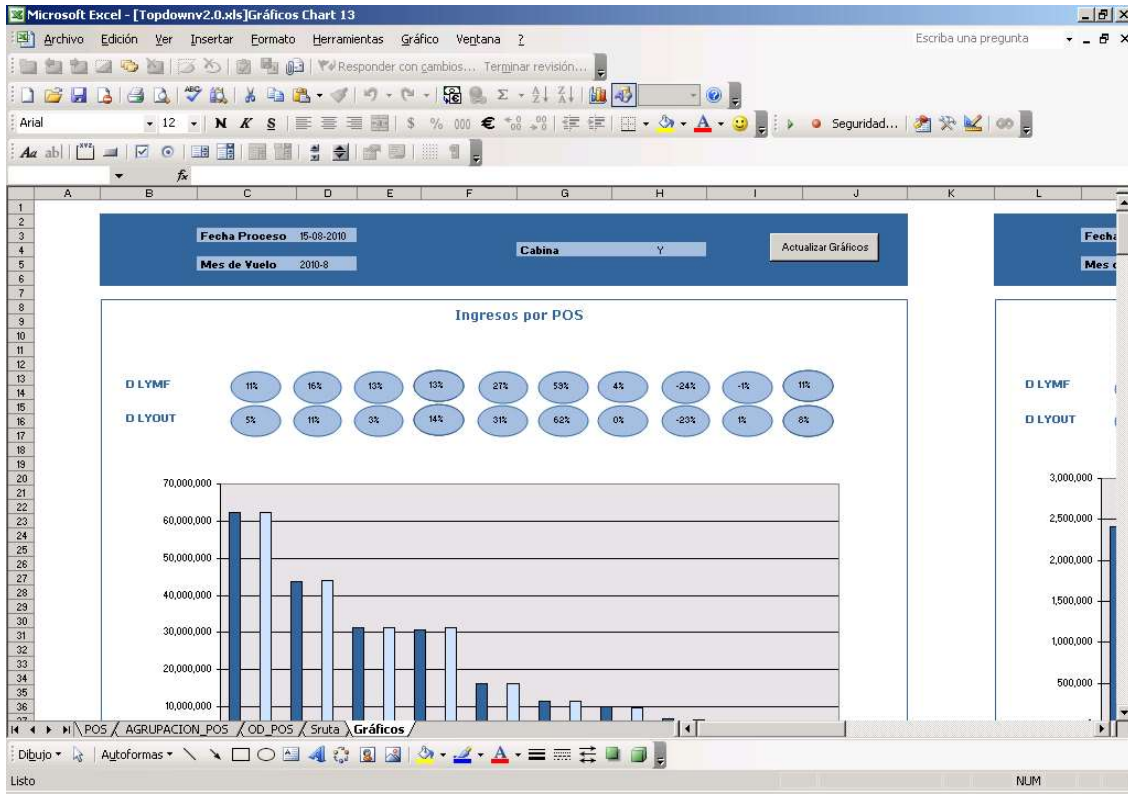


Figura 28: Vista gráficos

5.4.3 Arquitectura final del sistema

La forma propuesta para implementar el sistema tiene una ventaja relacionada a la flexibilidad, dado que pueden utilizarse herramientas/software independientes en cada una de las capas, por lo que fue posible elegir las herramientas idóneas para cada una de las partes del desarrollo.

El hecho de ocupar una solución como PALO (como motor OLAP) permite obtener la información agregada de forma rápida, ya que PALO es el gestor de las consultas dentro de Excel (al cambiar filtros, por ejemplo).

Para el resto de las capas, al realizar pruebas de procesamiento, se observó que los tiempos permiten procesar la información con todas las lógicas y unificarla para ser utilizada en el mismo día.

La arquitectura completa del sistema, capa de datos, lógica y visualización se describe a continuación:

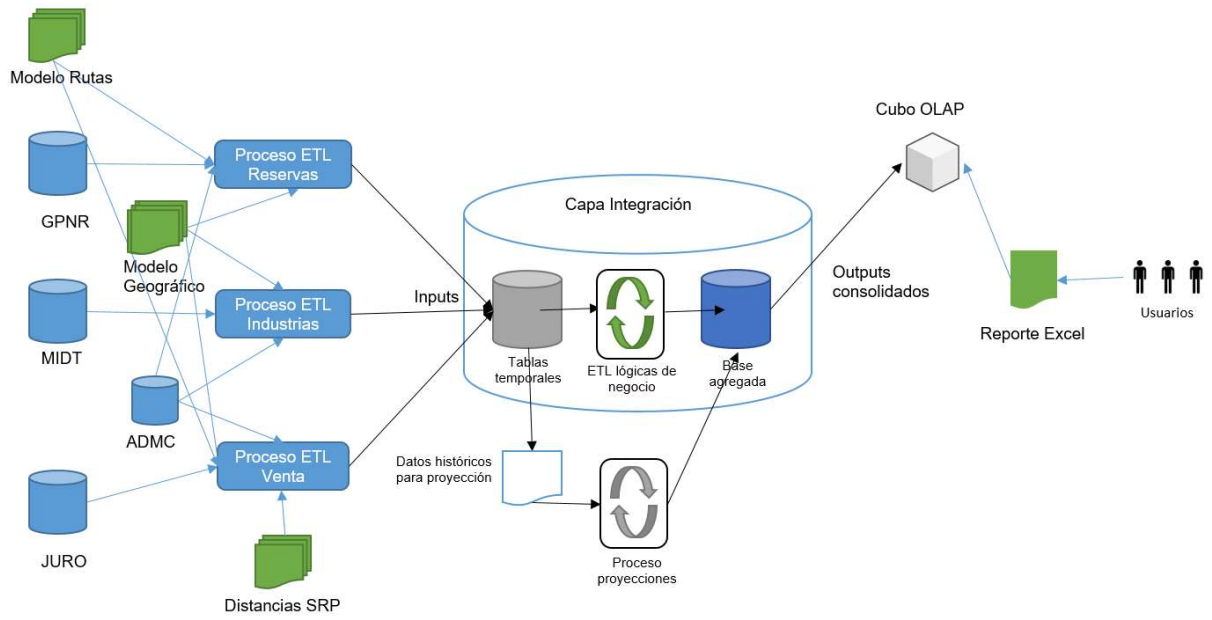


Figura 29: Modelo final

6. ANALISIS DE LOS RESULTADOS

Los resultados del proyecto lo podemos separar en:

- Eficiencia de la interfaz o herramienta y calidad de información
- Indicador de proyecciones como estimador de la demanda.
- Beneficios de la herramienta

Eficiencia de la herramienta

La herramienta desarrollada cumple con los objetivos del proyecto ya que presenta la información de forma clara y transversal para todos los usuarios, utilizando lógicas de negocio previamente consensuadas. Esto último permite que la comunicación entre los diferentes equipos sea más rápida y que la toma de decisiones se base en indicadores comunes para todos.

En términos de utilización de la herramienta, ésta presenta un tiempo de respuesta a consultas promedio entre 15 a 30 segundos. Para consultas más complejas, donde el despliegue de información es mayor, los tiempos de respuesta se encuentran alrededor de 1 minuto. Esto se debe a que una acción implica la ejecución en paralelo de múltiples consultas sobre un universo de más de 30 millones de registros.

Adicional a la rapidez en la lectura, la herramienta también permite a todos los usuarios concurrentes autorizados la modificación de los valores de proyección, realizando los ajustes que sean necesarios. Esta modificación es vista por otros los otros usuarios concurrentes.

Otro punto favorable de la herramienta es su tiempo de preparación. El procesamiento de la información previo a la generación del reporte tarda aproximadamente 4 horas, lo cual permite contar en el mismo día con la información más actualizada de las bases de datos de la compañía.

Indicador de proyecciones como estimador de la demanda

Los resultados de los modelos de predicción de demanda son favorables, ya que para aproximadamente el 90% de las agrupaciones de origen destino, el error se encuentra bajo el 10%, en ambientes simulados.

A medida que el horizonte de tiempo se amplía, el error es mayor. Es decir, mientras menos información se tenga con respecto al futuro, usualmente asociado a la anticipación del periodo de vuelo consultado, la proyección va decreciendo en exactitud.

Adicionalmente, se observa que las agrupaciones de origen destino que concentran una mayor cantidad de reservas presentan un error menor, en relación con agrupaciones con menor cantidad de reservas en los periodos observados. Esto es explicado debido a que

las agrupaciones origen destino que tienen mayor demanda representan mercados más estables del negocio y no sufren de variaciones atípicas.

Beneficios de la herramienta

Uno de los principales beneficios que aporta este tipo de herramientas a la organización es que es posible reducir el tiempo que los analistas de *Revenue Management* destinan a la recolección de datos y así enfocar los esfuerzos en realizar análisis adhoc que aporten a la mejor gestión de la ruta, con información consolidada para los procesos de negocio internos.

Una de las tareas más complejas realizadas por los analistas era la recolección y agregación de datos, lo cual podía tardar hasta más de un día. Sumado a esto, existía la complejidad que la extracción de información no se realizaba de manera transversal por los distintos equipos, generando diferencias al momento de querer consensuar alguna decisión de negocio a nivel transversal entre diferentes áreas.

Con la herramienta, las tareas de coordinación entre los equipos se enfocan en la gestión y toma de decisiones, no en discusiones e interrogantes relacionadas a la información, pues los cálculos para los indicadores se encuentran unificados y transparentados para todos los equipos. Así, se produce una reducción en los costos de coordinación entre distintas áreas de la compañía.

7. CONCLUSIONES Y COMENTARIOS FINALES

El escenario actual del negocio aéreo para la región se presenta como un momento desfavorable para los principales actores del continente. Esto se debe principalmente a la inestabilidad económica de los países latinoamericanos, y el largo periodo que se ha mantenido esta situación.

Si bien esto reduce las oportunidades de corto plazo en cuanto a crecimiento de la aerolínea en oferta y capacidad, si representa una oportunidad de aprendizaje y apertura de posibilidades a mediano plazo, ya que el escenario adverso hace poner foco en las actividades relacionadas con la optimización de costos y el rediseño de procesos.

Adicionalmente, dadas las características de la industria aérea, las prioridades hoy en día se centran en la reducción de costos y la simplificación de procesos, como forma de ser más flexibles y competitivos.

Teniendo en cuenta lo anterior, se identifica que una de las funciones de mayor incidencia en los resultados del negocio aéreo de pasajeros se encuentra en el área de *Revenue Management*. Este equipo gestiona un proceso de mejora continua que busca la maximización de ingresos, sujeto a las restricciones de capacidad de la flota que opera en las rutas que maneja. Es por esto que las empresas aéreas buscan estar a la vanguardia de esta disciplina, pues esto les permitirá obtener mayores. LATAM no se queda atrás y se encuentra en una constante búsqueda de oportunidades de mejora.

Dentro de *Revenue Management*, la gestión de información en búsqueda de *insights* y la proyección de demanda de pasajeros futura es relevante dentro del proceso táctico comercial diario, así como también para la gestión cruzada con otras áreas que visualizan oportunidades a futuro.

Sin embargo, en el trabajo de gestión de información realizado por los equipos de Revenue Management, se encuentran dos factores que pueden complejizar el trabajo realizado por esta área:

- La información se encuentra distribuida en diversas fuentes de información
- La forma de extraer la información no se encuentra estandarizada o consensuada con los equipos con los cuales se toman las decisiones.

Estos puntos tienen como consecuencia que los distintos equipos pueden extraer la información de fuentes y formas diferentes, generando diferencias en la toma de decisiones. El costo de coordinación aumenta y los procesos se hacen más complejos.

Es por esto que se realiza un trabajo de levantamiento de los indicadores requeridos, unificación de las diferentes reglas de negocio y la implementación de un modelo de proyección de demanda como apoyo a las proyecciones que ya trabajan los analistas.

Lo anterior se implementa mediante la construcción de un sistema de información que reúne las reglas de negocio, la información y las proyecciones para potenciar la búsqueda de oportunidades en el día a día.

Un aspecto importante de la implementación es contar con un desarrollo rápido, una herramienta que fuera potente, pero a la vez fácil de utilizar y visualmente familiar para los usuarios. Es por ello que se decide utilizar Microsoft Excel para la interfaz, apoyada del motor OLAP PALO para el manejo de grandes volúmenes de información, y del software Pentaho para la extracción, transformación y carga de data a las bases de datos que soportan la herramienta.

La complejidad del trabajo radica en la necesidad de mezclar diversas disciplinas, conocimiento de herramientas informáticas y el entendimiento tanto de los conceptos de negocio como de los sistemas internos de la aerolínea.

La aplicación práctica de modelos de proyección complejos puede tener resultados de poca precisión, dependiendo de la información que se le entregue al modelo. En las pruebas se observó que en algunas ocasiones los modelos más simples podían cumplir bien la tarea, y en otras, era necesaria la aplicación de un modelo más complejo. Es por esto que el desarrollo de la herramienta con un módulo adaptativo, en el cual automáticamente se elige el modelo con menor error, otorga buenos resultados.

Sin embargo, en algunos no era posible encontrar un modelo predictivo que tuviera un rango de error aceptable o que respetara las condiciones de negocio impuestas. Estos casos eran explicados por cambios bruscos en la demanda de los mercados, dados por variables externas (por ejemplo, eventos, condiciones meteorológicas entre otros). En estas ocasiones se establecía como proyección el valor de demanda del año pasado (LYMF) ajustado por un factor de diferencial de anticipación.

La incorporación de data adicional al reporte permite consolidar toda la información necesaria para la toma de decisiones. Cuando la información se encuentra distribuida en diversas fuentes es difícil su comparación y análisis, sin embargo, esta herramienta permite otorgar una mirada consolidada y una unificación de las versiones de los indicadores y sus análisis.

El sistema fue modelado por capas, creando varios sistemas más pequeños para llegar a un sistema de información consolidado. Esto permite un desarrollo más rápido que otorga beneficios mayores que desarrollar cada uno por separado: datos, lógica y visualización.

La herramienta propuesta en este proyecto es una solución novedosa e inexistente en la compañía para la época. La arquitectura planteada fue una innovación en cuanto al concepto de manejo de información en la empresa, permitiendo así el consenso para utilizar indicadores transversales a diferentes negocios, la coordinación de diferentes equipos basados en sistemas de información que disponibilizaran de forma periódica data relevante para la toma de decisiones, en un mismo lugar y con indicadores comparables.

En este trabajo, se cumplieron todos los objetivos planteados inicialmente, siendo posible también la implementación de un prototipo funcional de la herramienta.

7.1 Mejoras y trabajos futuros

7.1.1 Construcción de reglas automáticas complejas para el despliegue de alertas

Un complemento a la herramienta ya desarrollada es la implementación de alertas de, por ejemplo, tendencias crecientes, mercados caídos, anomalías en los mercados que podrían ser explicadas por eventos externos etc. Estas alertas pueden ser generadas con la información disponible, disminuyendo el tiempo destinado a la revisión de la información.

7.1.2 Incorporación de un módulo para incorporar influencias externas al modelo

En el negocio aéreo es frecuente ver que la cantidad de reservas de un mercado determinado, para una fecha determinada, se ve afectada por eventos externos, que pueden fomentar la demanda para ese mercado o bien contraerla. Estos eventos son puntuales, pero para mercados más pequeños pueden significar una gran alternación en las proyecciones realizadas. Por eso, se recomienda en el futuro incorporar un módulo que permita especificar en el sistema las fechas que no debieran ser consideradas para la proyección, debido a las alteraciones en la demanda generada por estos eventos. Así, se puede establecer una proyección propuesta igual a la demanda del año pasado ajustada por un cierto factor e indicar a los usuarios que el valor se ve afectado por eventos puntuales externos.

7.1.3 Visualización

En el futuro es posible revisar la visualización desarrollada para aumentar su usabilidad y compatibilidad con los procesos de negocio y sus usuarios. Por ejemplo, con el desarrollo de una versión web sustentada en PALO OLAP.

7.1.4 Construcción de un *Data Warehouse*

Si bien el tipo de soluciones como la herramienta desarrollada apunta a entregar un alto valor en un corto periodo de tiempo y a menor costo, a futuro se visualiza la necesidad de llevar toda la información y lógicas a un *Data Warehouse* donde se consoliden todas las fuentes e indicadores, con tiempos de respuesta y niveles de servicio acordes con el crecimiento de la compañía y sus necesidades de información.

7.1.5 Mejoras en la metodología de proyección

Si bien los resultados a nivel agrupado otorgan una visión correcta, con errores dentro del rango definido como aceptable, la explotación de mayor información del cliente y

técnicas más avanzadas lograrían mejores resultados para elementos desagregados, a mayor nivel de detalle.

La tendencia apunta a que los sistemas inteligentes como el desarrollado en esta memoria irán en aumento en todas las industrias, ya que manejan el activo más grande dentro de una empresa: la información.

8. BIBLIOGRAFIA

- [1] ATHANASIOS ZAKHARY, NEAMAT EL GAYAR, AMIR F. ATIYA. 2008. A Comparative Study of the Pickup Method and its Variations Using a Simulated Hotel Reservation Data. ICGST-AIML Journal, Volume 8, Special Issue on Computational Methods for the Tourism Industry.
- [2] ATHANASIOS ZAKHARY, AMIR F. ATIYA, HISHAM EL-SHISHINY. 2009. Forecasting Hotel Arrivals and Occupancy Using Monte Carlo Simulation.
- [3] CAPTAIN JOAN H. NEWBERNE. 2006. Holt-Winters Forecasting: A Study of Practical Applications for Healthcare.
- [4] HAN LIN SHANG, ROB J HYNDMAN. 2009. Non parametric time series forecasting with dynamic updating.
- [5] DEPARTAMENTO DE INGENIERIA CIVIL INDUSTRIAL, UNIVERSIDAD DE CHILE, 2008. Apuntes de Gestión de operaciones. IN47A-1. Universidad de Chile, Chile.
- [6] World Airlines awards, aerolíneas [en línea] <<http://www.worldairlineawards.com>> [consulta: marzo 2016].
- [7] Yahoo finance, indicadores LATAM Airlines [en línea] <<http://finance.yahoo.com>> [consulta: marzo 2016].
- [8] Wikipedia, Bases de datos [en línea] <https://es.wikipedia.org/wiki/Base_de_datos> [consulta: marzo 2016].
- [9] Wikipedia, ETL [en línea] <https://es.wikipedia.org/wiki/Extract,_transform_and_load> [consulta: marzo 2016].
- [10] Sabre, comunicado [en línea] <<https://www.sabre.com/insights/releases/latam-airlines-group-and-its-affiliated-carriers-select-sabre-technology-to-unify-and-power-its-passenger-reservations/>> [consulta: marzo 2016].
- [11] ÁREA DE MEJORES PRACTICAS REVENUE MANAGEMENT, LAN AIRLINES. 2012. Apuntes de Curso FARM. Chile, LAN Airlines.
- [12] SANCHIS, RAQUEL; POLER, RAÚL, ORTIZ, ÁNGEL. 2009. Técnicas para el Modelado de Procesos de Negocio en Cadenas de Suministro. Inf. tecnol. [online]. vol.20, n.2

9. ANEXOS

A. Software proyección demanda

Para aplicar un proceso recurrente de proyección se desarrolló en lenguaje java un robot de proyecciones de demanda, se utilizó además clases pre definidas de carácter opensource, *Openforecast* (<http://openforecast.sourceforge.net/>), que sirvió para estructurar las clases básicas del robot.

Como muestra dejo expresado el código de la clase main para el robot de gestión PNR

```
import java.io.BufferedReader;
import java.io.BufferedWriter;
import java.io.File;
import java.io.FileReader;
import java.io.FileWriter;
import java.io.IOException;
import java.io.PrintWriter;
import java.util.Date;
import java.util.Iterator;

import net.sourceforge.openforecast.DataPoint;
import net.sourceforge.openforecast.DataSet;
import net.sourceforge.openforecast.Forecaster;
import net.sourceforge.openforecast.ForecastingModel;
import net.sourceforge.openforecast.Observation;
import net.sourceforge.openforecast.models.MovingAverageModel;
import net.sourceforge.openforecast.models.NaiveForecastingModel;
import net.sourceforge.openforecast.models.PolynomialRegressionModel;
import net.sourceforge.openforecast.models.RegressionModel;
```

```

import net.sourceforge.openforecast.models.SimpleExponentialSmoothingModel;
import net.sourceforge.openforecast.models.DoubleExponentialSmoothingModel;
import net.sourceforge.openforecast.models.TripleExponentialSmoothingModel;
import net.sourceforge.openforecast.output.DelimitedTextOutputter;

    public static void main(String[] args)
    {

        for(int s=1; s<=6; s++){

            System.out.println("Analizando input"+s+"");

                String aux;

                int aux2;

                String aux3;

                String aux4="";

                int contador=0;

                int aux25;

                String input="C:\\Forecast\\Input\\Transformaciones OUT\\GPNR
OUT\\input"+s+".txt";

                String
output_tes="C:\\Forecast\\Output\\GPNR_out\\outputtes"+s+".csv";

                String
output_reg="C:\\Forecast\\Output\\GPNR_out\\outputreg"+s+".csv";

                String
output_ses="C:\\Forecast\\Output\\GPNR_out\\outputses"+s+".csv";

                String
Resultado_tes="C:\\Forecast\\Resultados\\GPNR_out\\resultadotes"+s+".txt"+"";

                String
Resultado_reg="C:\\Forecast\\Resultados\\GPNR_out\\resultadoreg"+s+".txt"+"";

```

```

        try {BufferedReader in2 = new BufferedReader (new FileReader(new
File(input)));

        while(in2.readLine() !=null){

        contador=contador+1;}

in2.close();} catch(IOException in2){}

int auxiliar=0;

int i2=0;

int serie=0;

double marca=-99;

double[][] datotest=new double[(contador)/4][50];

try {

BufferedReader in = new BufferedReader (new FileReader(new
File(input)));

in.readLine(); // primera linea

for(int i=1;i<=contador-1;++i)

{

aux = in.readLine();

aux2=aux.indexOf(";");

aux25=aux.length();

aux3=aux.substring(0,aux2);

aux4=aux.substring(aux2+1, aux25);

```

```

serie=Integer.parseInt(aux3.trim());

    if (serie-1==auxiliar){
        if (i>1){datotest[serie-1][i2]=Double.parseDouble(aux4.trim());
i2=i2+1;}
        else{datotest[0][0]=Double.parseDouble(aux4.trim()); i2=1;}}
        else{

            datotest[serie-2][i2]=marca;
            datotest[serie-1][0]=Double.parseDouble(aux4.trim());
            i2=1;
            auxiliar=auxiliar+1;}

        if (i==contador-1){datotest[serie-1][i2]=marca;}

    }

    in.close();

        } catch (IOException in) {

            System.out.println("Error en la lectura del archivo"+"
"+in.getMessage());

        }

        DataSet initDataSet=new DataSet();

        DataSet[] prueba= new DataSet[contador];

        DataPoint dp;

```

```

int y=0;

for(int i=0;i<=serie-1;++i){
    DataSet initDataSet2=new DataSet();
    for(int z=0; datotest[i][z]!=marca; z++){
        dp=new Observation(datotest[i][z]);
        dp.setIndependentValue("t", z );
        initDataSet2.add(dp);
        y++;
        prueba[i]=initDataSet2;
    }
}

for(int i=0; i<=serie-1; i++){

    if(
        datotest[i][0]!=0 &&    datotest[i][1]!=0 && datotest[i][2]!=0
        && datotest[i][3]!=0 &&  datotest[i][4]!=0 && datotest[i][5]!=0
        && datotest[i][6]!=0 &&  datotest[i][7]!=0 && datotest[i][8]!=0
        && datotest[i][9]!=0 &&  datotest[i][10]!=0 && datotest[i][11]!=0
        && datotest[i][12]!=0 && datotest[i][13]!=0 && datotest[i][14]!=0
        && datotest[i][15]!=0 && datotest[i][16]!=0 && datotest[i][17]!=0
        && datotest[i][18]!=0 && datotest[i][19]!=0 && datotest[i][20]!=0
        && datotest[i][21]!=0 && datotest[i][22]!=0 && datotest[i][23]!=0
        && datotest[i][24]!=0){

```

```

prueba[i].setTimeVariable("t");
prueba[i].setPeriodsPerYear(12);

// Get "best fit" triple exponential smoothing model
ForecastingModel tesModel
=
TripleExponentialSmoothingModel.getBestFitModel(prueba[i]);
DataSet ResultadoExpoTriple= getDataSet(prueba[i], 12);
tesModel.forecast(ResultadoExpoTriple);
//System.out.println("Proyectando serie
"+ResultadoExpoTriple.toString()+"");
try{
DelimitedTextOutputter out3=new
DelimitedTextOutputter(output_tes, true);
out3.output(ResultadoExpoTriple);
out3.close();
}catch (IOException out3){System.out.println("Error en la
escritura del archivo CSV"+" "+out3.getMessage());}

try{ PrintWriter archivojoin= new PrintWriter(new
BufferedWriter(new FileWriter(Resultado_tes, true)));
archivojoin.println(i+1);
archivojoin.close();}catch (IOException out3){}

} // cierra if

```



```

        else{ if( datotest[i][24]!=0 && datotest[i][23]!=0 &&
datotest[i][22]!=0 && datotest[i][21]!=0

                && datotest[i][0]!=0 && datotest[i][6]!=0

                && datotest[i][1]!=0 && datotest[i][2]!=0

                && datotest[i][3]!=0 && datotest[i][4]!=0 &&
datotest[i][5]!=0

                && datotest[i][6]!=0 && datotest[i][7]!=0 &&
datotest[i][8]!=0

                && datotest[i][9]!=0 && datotest[i][10]!=0 &&
datotest[i][11]!=0

                ){

prueba[i].setTimeVariable("t");

prueba[i].setPeriodsPerYear(12);

ForecastingModel RegModel
=new RegressionModel("t");

RegModel.init(prueba[i]);

DataSet ResultadoRegModel=getDataSet(prueba[i],12);

RegModel.forecast(ResultadoRegModel);

//System.out.println("Proyectando serie
"+ResultadoRegModel.toString()+"");

        try{DelimitedTextOutputter outreg=new
DelimitedTextOutputter(output_reg, true);

                outreg.output(ResultadoRegModel);

                outreg.close();

                }catch (IOException outreg){System.out.println("Error en la
escritura del archivo CSV"+" "+outreg.getMessage());}

```

```

        ForecastingModel desModel
        = DoubleExponentialSmoothingModel.getBestFitModel(prueba[i]);
        DataSet ResultadoExpoDoble = getDataSet(prueba[i],12);
        desModel.forecast(ResultadoExpoDoble);

        try{DelimitedTextOutputter out2=new
DelimitedTextOutputter(output_ses, true);

        out2.output(ResultadoExpoDoble);

        out2.close();

        }catch (IOException out2){System.out.println("Error en la
escritura del archivo CSV"+" "+out2.getMessage());};

        try{ PrintWriter archivojoin= new PrintWriter(new
BufferedWriter(new FileWriter(Resultado_reg, true)));

        archivojoin.println(i+1);

        archivojoin.close();}catch (IOException out3){}

    }// Cierre IF

    }//cierre else 2do if

    }// Cierre del FOR

} //cierre "for" semanas
}
/**

```

```

    * Funcion que crea espacio en el DataSet para el pronostico (from
startIndex to

    * endIndex) of object into an OpenForecast DataSet.

    * @param series el dataset al cual modificare para el pronostico.

    * @param startIndex el index del primer mes que se pronosticara.

    * @param endIndex el index del ultimo elemento que se pronosticara.

    * @return Un OpenForecast DataSet representando los puntos para
pronosticar.

    */

    public static DataSet getDataSet( DataSet series,

                                    int endIndex )

    {

        DataSet auxi=new DataSet();

        int aux=series.size();

        for ( int i=aux; i<(aux+endIndex); i++ )

        {

            DataPoint dp = new Observation(0.0);

            dp.setIndependentValue("t", i );

            auxi.add( dp );

        }

        return auxi;

    }

    public DataSet getForecastDataSet( ForecastingModel model,

                                       DataSet initDataSet,

                                       int endIndex)

    {

```

```

// Initialize the forecasting model
model.init( initDataSet );

// Get range of data required for forecast
DataSet fcDataSet = getDataSet( initDataSet, endIndex );

// Obtain forecast values for the forecast data set
model.forecast( fcDataSet );

//Obtener descripcion del modelo y el error
System.out.println(model.toString());
System.out.println("MAD =" +model.getMAD()+"");
System.out.println("MAPE =" +model.getMAPE()+"");
System.out.println("MSE =" +model.getMSE()+"");
System.out.println("SAE =" +model.getSAE()+"");
System.out.println("Bias =" +model.getBias()+"");

return fcDataSet;
}
}

```

B. Pentaho data integration

Información de pentaho: http://www.pentaho.com/products/data_integration/

Las transformaciones en *Pentaho Data integration* modifican flujos de datos para transformarlos en información, para este proyecto fueron realizadas decenas de transformación que se ejecutan semanalmente (de forma automática):

Un caso de ejemplo para la extracción de datos y prepararlos para la proyección:

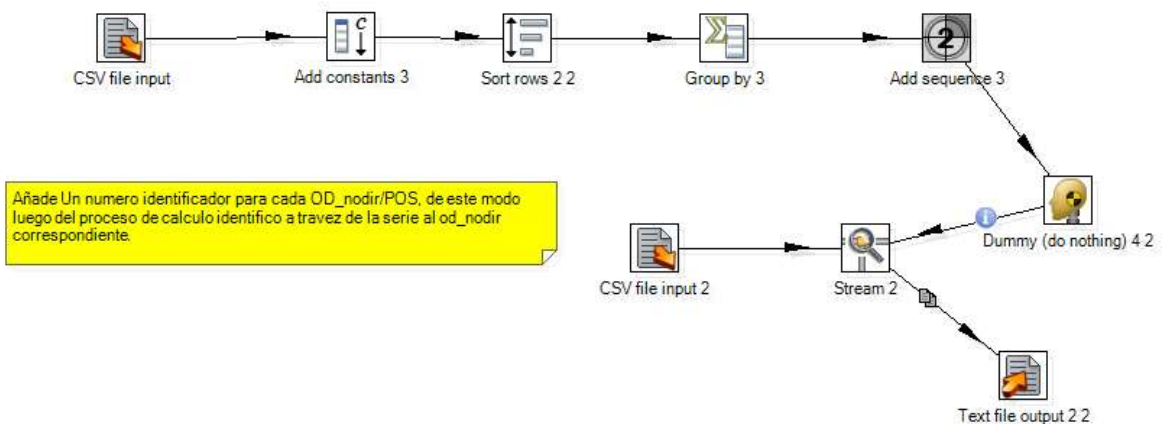
1. Ordenamiento de PNRs:

Se ejecuta para ordenar los pnr's de forma cronológica para que entren dentro de los modelos de proyecciones



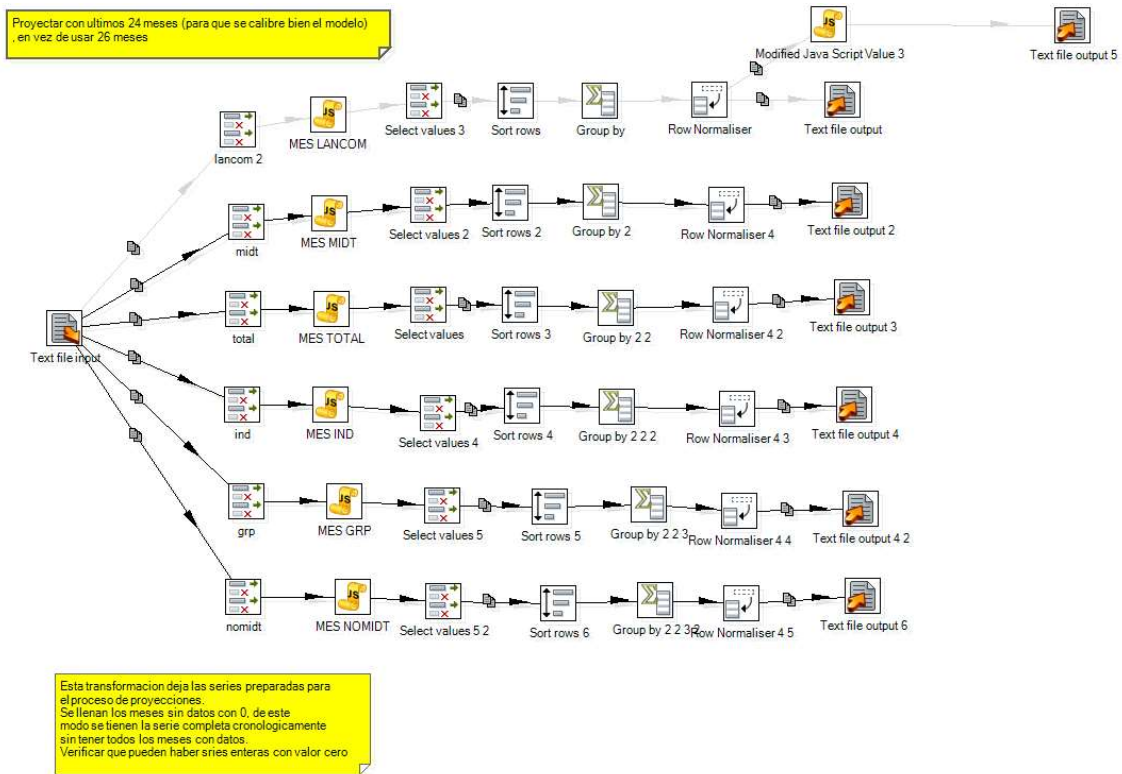
2. Añade serie a cada ODPOS.

Añade números identificadores a cada origen-destino POS, para después identificarlo en el proceso de proyecciones.



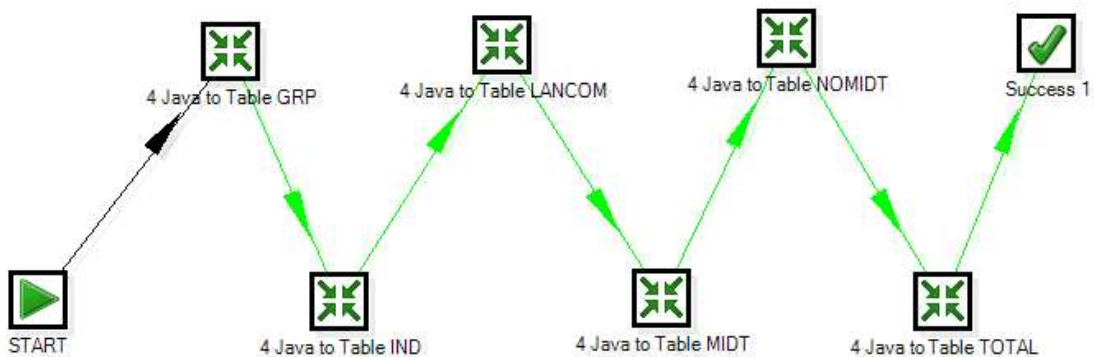
3. Transformación que construye la estructura final para ser proyectados

Se separa el flujo total de datos previo de acuerdo a la lógica del negocio, una vez que esta transformación termina se tienen los archivos input del programa que proyecta la demanda.

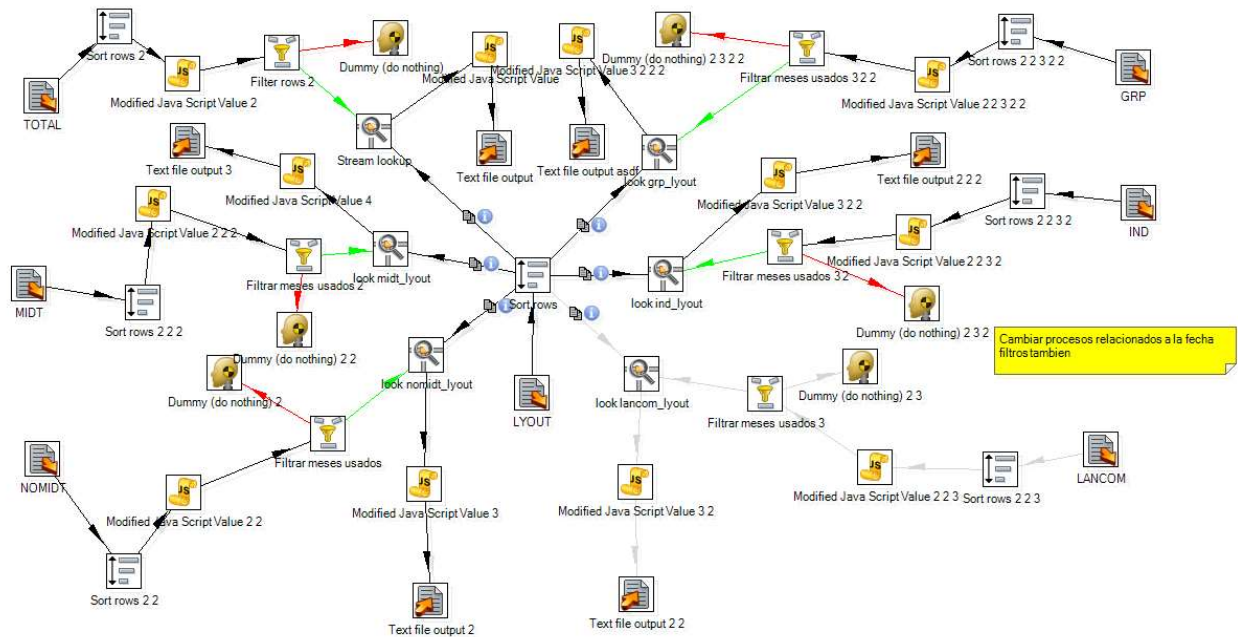


4 .Jobs de java a txt

Este Job (o agrupación de transformaciones) transforma la información desde el proceso de proyecciones de vuelta a las bases de datos (Capa de integración)



5. Transformación a agrupar datos totales de la proyección



C. PALO

Es una herramienta de BI de la empresa JEDOX, uno de sus productos es un *plugin* asociado a Excel y que permite generar cubos OLAP que se almacenan en memoria en un servidor independiente, lo que ayuda enormemente a la rapidez al momento de consultar.

Align your Excel data

Palo for Excel

Does this sound familiar? Reporting and Budget Planning are handled with Excel and as a result, hundreds of sheets with often inconsistent data are circulating throughout the company?

Palo for Excel stops the Excel-Chaos. Palo saves your Excel data in a central business model. The data is available to all users via your network, quickly and easily.