



Herramienta de analítica visual para el manejo de irregularidades de equipajes

Emilio Andrés Alvarado Santis

Universidad de los Andes
Departamento de Ingeniería de Sistemas y Computación
Bogotá, Colombia
2020

Herramienta de analítica visual para el manejo de irregularidades de equipajes

Emilio Andrés Alvarado Santis

Trabajo de tesis presentado como requisito para optar al título de:
Magíster en Ingeniería de Información

Director:
Ph.D. José Tiberio Hernández Peñalosa

Línea de Investigación:
Visual Analytics
Grupo de Investigación:
IMAGINE-Computación Visual, I+D+i

Universidad de los Andes
Departamento de Ingeniería de Sistemas y Computación
Bogotá, Colombia
2020

A Dios, por estar siempre presente.

A mi esposa, por sus valiosos consejos.

A mis padres, por su incondicional apoyo.

Agradecimientos

A la Universidad de los Andes y su reconocido equipo de profesores, quienes aportaron sus valiosos conocimientos y experiencia en mi proceso de formación.

Al profesor José Tiberio Hernández, por la asesoría y acompañamiento durante el planteamiento y realización de este proyecto.

A Rosmery Bedoya y Edgar Otálora, por el aporte en el entendimiento de los procesos de negocio y sus necesidades. A Andrés Hurtado, por el acompañamiento inicial a este trabajo con su iniciativa de *flight delays* y por el apoyo técnico en las tareas de *data wrangling*. A Diego Díaz y César Avila, por su aporte en el entendimiento de los datos. A Lenin Hernández, por su apertura para el desarrollo de estas iniciativas dentro del equipo de trabajo. A Avianca, por la oportunidad para conocer el proceso de negocio y plantear esta propuesta.

Resumen

El manejo de equipajes hace parte integral de los procesos de la industria de transporte aéreo de pasajeros y juega un papel fundamental en la percepción de la calidad del servicio. El incremento continuo en el número de personas transportadas cada año conlleva desafíos importantes en la reducción de las irregularidades presentadas. La información generada en cada etapa del proceso es un insumo fundamental para habilitar capacidades de análisis que permitan su medición y mejora. Este trabajo identifica las principales necesidades en el manejo de irregularidades de equipajes, las fuentes de información involucradas y los métodos de análisis para abordar el problema; también se diseña e implementa una herramienta de analítica visual para apoyar las capacidades de análisis descriptivo y diagnóstico. El diseño incluye diferentes vistas para apoyar el análisis de datos de irregularidades de equipajes, vuelos, demoras, condiciones climáticas, vuelos de otras aerolíneas y uso de la infraestructura aeroportuaria; adicionalmente se generan datos para simular el comportamiento de las entradas y salidas de vuelos, así como la ubicación de las maletas en el contexto del aeropuerto en un momento dado. La solución es aplicada al caso de estudio conformado por los vuelos ubicados en la terminal 1 del Aeropuerto Internacional El Dorado, en la ciudad de Bogotá, en el periodo 2018-2019, para una aerolínea específica. Se realizan pruebas a 22 usuarios que ejecutan tareas de análisis de complejidad baja, media y alta, con niveles de acierto de 46%, 41% y 31%, respectivamente.

Palabras clave: analítica visual, manejo de irregularidades de equipajes.

Abstract

Baggage handling is an integral part of the passenger air transport industry processes and plays a fundamental role in the perception of service quality. The continuous increase in the number of people transported each year brings significant challenges in reducing the irregularities presented. The information generated at each stage of the process is a fundamental input to enable analysis capabilities that allow measurement and improvement. This work identifies the main needs in baggage handling irregularities, the sources of information involved and the analysis methods to address the problem. A visual analytics tool is also designed and implemented to support descriptive and diagnostic analysis capabilities. The design includes different views to support data analysis of baggage irregularities, flights, delays, weather conditions, flights by other airlines, and use of airport infrastructure. Additionally, data is generated to simulate the behavior of flight arrivals and departures, as well as the location of bags at the airport on a specific day. The solution is applied to the case study made up of flights located in Terminal 1 of El Dorado International Airport, Bogotá, in the period 2018-2019, for a specific airline. Tests are carried out on 22 users who carry out analysis tasks of low, medium and high complexity, with success levels of 46%, 41% and 31%, respectively.

Keywords: visual analytics, mishandled baggage handling.

Contenido

	Pág.
CAPÍTULO 1. MARCO TEÓRICO	3
1.1. CONTEXTO DE LA INDUSTRIA	3
1.1.1. <i>Cifras del mercado global</i>	3
1.1.2. <i>Proceso general: The baggage journey</i>	5
1.1.3. <i>Puntos de registro obligatorios</i>	6
1.1.4. <i>Indicadores del proceso</i>	8
1.2. ALTERNATIVAS PARA EL MEJORAMIENTO DEL PROCESO DE EQUIPAJES.....	8
1.2.1. <i>Modelamiento, optimización y simulación</i>	9
1.2.2. <i>Análisis de información espacio-temporal y en movimiento</i>	13
1.2.3. <i>Sistemas de manejo de equipajes</i>	16
1.2.4. <i>Herramientas para diseño de soluciones de analítica e inteligencia de negocios</i>	17
CAPÍTULO 2. PLANTEAMIENTO DEL PROBLEMA Y OBJETIVOS.....	23
2.1. PLANTEAMIENTO DEL PROBLEMA	23
2.2. OBJETIVO GENERAL	23
2.3. OBJETIVOS ESPECÍFICOS.....	23
CAPÍTULO 3. METODOLOGÍA	25
3.1. CAPACIDADES ANALÍTICAS	25
3.2. METODOLOGÍA ASUM-DM	26
3.3. MARCOS DE REFERENCIA PARA ANALÍTICA VISUAL	27
3.4. ABORDAJE DEL PROYECTO.....	29
CAPÍTULO 4. RESULTADOS	31
4.1. FASE 1: ANÁLISIS DESCRIPTIVO	31
4.1.1. <i>Entendimiento del Negocio</i>	31
4.1.2. <i>Aproximación analítica</i>	36
4.1.3. <i>Procesamiento de Datos</i>	36
4.1.4. <i>Modelamiento y Visualización</i>	39
4.1.4.1. <i>Diseño de la herramienta</i>	39
4.1.4.2. <i>Análisis asistido</i>	41
4.1.4.3. <i>Diseño de la visualización</i>	42
4.1.5. <i>Evaluación</i>	52
4.1.5.1. <i>Hallazgos relevantes en datos</i>	52
4.1.5.2. <i>Hallazgos en el uso de modismos de visualización</i>	52
4.1.6. <i>Despliegue</i>	54

4.1.7. <i>Retroalimentación</i>	54
4.1.7.1. <i>Pruebas funcionales</i>	54
4.1.7.2. <i>Pruebas de desempeño</i>	54
4.2. FASE 2: ANÁLISIS DIAGNÓSTICO.....	56
4.2.1. <i>Entendimiento del Negocio</i>	56
4.2.2. <i>Aproximación analítica</i>	59
4.2.3. <i>Procesamiento de Datos</i>	60
4.2.3.1. <i>Modelo conceptual de datos</i>	63
4.2.3.2. <i>Calidad y preparación de datos</i>	68
4.2.3.3. <i>Volumen de datos</i>	69
4.2.3.4. <i>Entendimiento de los datos</i>	70
4.2.4. <i>Modelamiento y Visualización</i>	71
4.2.4.1. <i>Diseño de la herramienta</i>	71
4.2.4.2. <i>Análisis asistido</i>	72
4.2.4.3. <i>Diseño de la visualización</i>	73
4.2.5. <i>Evaluación</i>	99
4.2.5.1. <i>Hallazgos relevantes en datos</i>	99
4.2.5.2. <i>Hallazgos en el uso de modismos de visualización</i>	100
4.2.6. <i>Despliegue</i>	101
4.2.7. <i>Retroalimentación</i>	101
4.2.7.1. <i>Pruebas funcionales</i>	101
4.2.7.2. <i>Pruebas de desempeño</i>	103
4.3. ANÁLISIS Y DISCUSIÓN	105
CAPÍTULO 5. CONCLUSIONES Y RECOMENDACIONES.....	111
5.1. CONCLUSIONES.....	111
5.2. RECOMENDACIONES Y TRABAJO FUTURO	112

Lista de figuras

	Pág.
Figura 1. Tendencia de pasajeros transportados e incidentes de equipajes 2007-2017.....	3
Figura 2. Causas de equipaje retrasado e incidentes presentados en 2017.	4
Figura 3. Ejemplo de Códigos RL.	5
Figura 4. <i>The baggage journey</i>	6
Figura 5. Puntos de registro obligatorios en <i>The Baggage Journey</i>	7
Figura 6. Un sistema BHS y sus componentes.	10
Figura 7. Infraestructura para el subprocesso de equipaje entrante.....	11
Figura 8. Puntos y tiempos en el subprocesso de manejo de equipaje entrante.....	11
Figura 9. Vehículos con código de destino (DCV).....	12
Figura 10. Tipología de tareas para análisis de datos espaciotemporales.	13
Figura 11. Cubo de espacio-tiempo.	14
Figura 12. Trayectoria de una pieza de equipaje dentro del sistema BHS.	15
Figura 13. Marco de trabajo para el análisis de trayectorias.	15
Figura 14. Visualización de puntos de congestión.	16
Figura 15. Ejemplo de sistema BHS.	17
Figura 16. Cuadrante de Gartner para plataformas de BI y analítica.....	18
Figura 17. Las 4 capacidades analíticas.	25
Figura 18. Metodología ASUM-DM.	26
Figura 19. Proceso de analítica visual.	28
Figura 20. Framework para el diseño de visualizaciones.	29
Figura 21. Abstracción del flujo de gestión de equipaje	31
Figura 22. Flujo de gestión de una irregularidad	34
Figura 23. Modelo de datos en el análisis descriptivo	37
Figura 24. Vista lógica de la solución	40
Figura 25. Vista física de la solución	40
Figura 26. Análisis asistido - correlación entre atributos numéricos	42
Figura 27. Vista general de la solución para el análisis descriptivo	49
Figura 28. Coordinación de vistas (cifras por aeropuerto para la región Europa)	50
Figura 29. Coordinación de vistas (saqueos en Norteamérica)	51
Figura 30. Alineación de diagramas de tiempo	51
Figura 31. Coordenadas paralelas	53
Figura 32. Mapa de aeropuertos responsables, irregularidades e indicador MB	53
Figura 33. Mapa mental de información de contexto de los equipajes	58

Figura 34. Variables de análisis y disponibilidad de datos	62
Figura 35. Modelo conceptual de datos en el análisis diagnóstico	63
Figura 36. Vista física de la solución – análisis diagnóstico	72
Figura 37. Correlación entre atributos numéricos en análisis diagnóstico	72
Figura 38. Vista de Vuelos y Demoras	80
Figura 39. Vista de Condiciones Climáticas	83
Figura 40. Vista de Contexto	87
Figura 41. Vista de Contexto - Vuelos ingresando (entradas) a El Dorado.....	90
Figura 42. Vista de Contexto - Vuelos saliendo de El Dorado	94
Figura 43. Guía para el usuario	98
Figura 44. Resultados de las pruebas con usuarios.....	102

Lista de tablas

	Pág.
Tabla 1. Reporte MB para aerolíneas americanas a octubre de 2018.....	8
Tabla 2. Estructura iterativa del plan de trabajo a alto nivel basado en la metodología....	30
Tabla 3. Descripción del flujo de gestión de equipajes.....	32
Tabla 4. Descripción del flujo de gestión de una irregularidad	35
Tabla 5. Conjunto de datos de irregularidades.....	37
Tabla 6. Conjunto de datos de vuelos	38
Tabla 7. Volumen de datos	39
Tabla 8. Ejemplo de hallazgos del análisis descriptivo asistido para 2018	41
Tabla 9. Tipos de atributos de las irregularidades.....	43
Tabla 10. Tipos de atributos de los vuelos	43
Tabla 11. Decisiones de diseño para la visualización	46
Tabla 12. Ejemplo de mediciones de tiempo en las interacciones.....	55
Tabla 13. Sistemas involucrados en la gestión de equipajes	59
Tabla 14. Fuentes de información adicionales	60
Tabla 15. Conjunto de datos de vuelos con demoras.....	64
Tabla 16. Conjunto de datos de aeropuertos	65
Tabla 17. Conjunto de datos de clima.....	65
Tabla 18. Conjunto de datos de aeronaves.....	65
Tabla 19. Conjunto de datos de aerolíneas.....	65
Tabla 20. Conjunto de datos de estadísticas de El Dorado	66
Tabla 21. Conjunto de datos de estadísticas de Aerocivil	66
Tabla 22. Conjunto de datos de estadísticas de FlightStats	66
Tabla 23. Conjunto de datos de zonas.....	67
Tabla 24. Conjunto de datos de ubicaciones	67
Tabla 25. Conjunto de datos de tiempos de tránsito	67
Tabla 26. Conjunto de datos de seguimiento de maletas.....	67
Tabla 27. Conjunto de datos de maletas.....	68
Tabla 28. Volumen de datos del análisis diagnóstico	69
Tabla 29. Interpretación de los valores para reglas de vuelo y visibilidad	70
Tabla 30. Interpretación de los valores para el atributo nubosidad.....	70
Tabla 31. Tipos de atributos de los vuelos y demoras.....	73
Tabla 32. Tipos de atributos de aeropuertos	73

Tabla 33. Tipos de atributos del clima.....	74
Tabla 34. Tipos de atributos de las aeronaves.....	74
Tabla 35. Tipos de atributos de aerolíneas	74
Tabla 36. Tipos de atributos para las estadísticas de El Dorado.....	75
Tabla 37. Tipos de atributos para las estadísticas de Aerocivil	75
Tabla 38. Tipos de atributos para las estadísticas de FlightStats	75
Tabla 39. Tipos de atributos para las Zonas	76
Tabla 40. Tipos de atributos para las Ubicaciones	76
Tabla 41. Tipos de atributos para el seguimiento de maletas.....	76
Tabla 42. Tipos de atributos para las maletas.....	77
Tabla 43. Decisiones de diseño para la vista de vuelos y demoras.....	81
Tabla 44. Decisiones de diseño para la vista de condiciones climáticas	84
Tabla 45. Decisiones de diseño para la vista de contexto.....	88
Tabla 46. Decisiones de diseño para la vista de entradas a El Dorado.....	91
Tabla 47. Decisiones de diseño para la vista de salidas	95
Tabla 48. Mediciones de tiempo en las interacciones para la fase diagnóstica	104
Tabla 49. Modismos de visualización no convencionales para PowerBi	113

Listado de términos y abreviaturas

A continuación se presentan los términos y acrónimos que aparecen referenciados a lo largo del documento:

- **Aerocivil:** La *Unidad Administrativa Especial de Aeronáutica Civil* o Aerocivil, es el organismo estatal colombiano encargado de vigilar y controlar el espacio aéreo y establecer políticas que permitan garantizar el correcto funcionamiento de la aviación civil bajo principios de seguridad, oportunidad y eficiencia.
- **BHS:** *Baggage Handling System* o sistema de manejo de equipajes, es un sistema de transporte instalado en los aeropuertos que conduce el equipaje desde los puntos de recepción (*counters*) hacia las áreas en donde son almacenados y posteriormente cargados a los aviones, en vuelos de salida. El sistema también transporta el equipaje que viene en los vuelos de llegada hacia los puntos de reclamo del equipaje por parte de los pasajeros o hacia un área de almacenamiento para ser cargado a otro vuelo, cuando se presentan conexiones.
- **BRS:** *Baggage Reconciliation System* o sistema de reconciliación de equipaje, es un sistema de seguimiento que registra el estado y ubicación del equipaje en todo su recorrido desde el counter de recepción hasta la aeronave (Amadeus, 2019c). Su objetivo es brindar información del equipaje en todo momento y reducir al mínimo las pérdidas.
- **IATA:** *International Air Transport Association* o Asociación Internacional de Transporte Aéreo, es una asociación conformada por 290 aerolíneas cuya función es formular estándares y políticas, así como prestar diferentes servicios para hacer del transporte aéreo una industria sostenible y segura (IATA, 2019b).
- **ICAO:** *International Civil Aviation Organization* o la *Organización de Aviación Civil Internacional (OACI)* es una agencia de la Organización de las Naciones Unidas creada en 1944 por el Convenio sobre Aviación Civil Internacional para estudiar los problemas de la aviación civil internacional y promover los reglamentos y normas únicos en la aeronáutica mundial (Wikipedia, 2020).

- **Make-Up Area:** es el área donde se clasifican las maletas y se asignan al vuelo correspondiente, antes de ser cargadas al tractor y luego a la aeronave.
- **Mishandled Baggage (MB):** una maleta mal manejada, según la traducción literal, es un reporte de un retraso, daño o robo de una maleta, la cual es registrada por una aerolínea o sus compañías intermediarias, en nombre del pasajero y que se gestiona como una reclamación (SITA, 2018). En este documento y en la solución implementada se utilizará el término '**incidente de equipaje**', **irregularidad o caso**, con el mismo significado.
- **NPS:** *Net Promoter Score*, indicador de satisfacción del cliente, mide la probabilidad de que el servicio recibido sea recomendado a otros clientes.
- **RL:** *Reason for Lost*, es una codificación de las causas por las cuales se presenta un incidente de equipaje.
- **SITA:** es una compañía multinacional que provee servicios de tecnología de información y comunicaciones a la industria de transporte aéreo. Sus principales clientes son aeropuertos y aerolíneas, y sus servicios están enmarcados en gestión comercial de aerolíneas, operación de pasajeros, vuelos y aeronaves, comunicaciones aire-tierra, gestión y operación de aeropuertos, operaciones de equipaje, seguridad en el transporte, gestión de carga, entre otros (SITA, 2019a).

Introducción

En la industria de transporte aéreo de pasajeros existen diferentes procesos y áreas con necesidades particulares en la búsqueda de obtener mejores niveles de satisfacción del servicio recibido por parte de los clientes, así como menores costos de operación de aerolíneas y aeropuertos. Con un volumen cada vez mayor de pasajeros transportados, se incrementan los retos en el cumplimiento de los itinerarios programados, en la reducción de las demoras en los vuelos y de los pagos por compensación a los clientes, ya sea por causa de afectación en el vuelo o por irregularidades presentadas con el equipaje.

Uno de los procesos de interés en esta industria es la gestión equipajes, el cual sumó 22.7 millones de irregularidades en el mercado global para el año 2017, asociadas a 4.08 billones de pasajeros transportados en el mismo año (SITA, 2018). Estas irregularidades representan miles de millones de dólares en pagos por compensación a pasajeros solamente por incidentes en equipajes, relacionados con una de las siguientes razones: demora, pérdida, daño o saqueo.

Los sistemas que actualmente se utilizan en aeropuertos y aerolíneas se enfocan en etapas específicas del proceso y carecen de capacidades que permitan analizar su comportamiento de manera integral: desde la entrega del equipaje a la aerolínea hasta la devolución al pasajero en el aeropuerto de destino. Lo anterior ha motivado la adquisición e integración de soluciones de tecnología específicas para las necesidades de cada compañía, típicamente a través de plataformas de inteligencia de negocios, bodegas de datos, reportes y tableros de control, que permiten apoyar las capacidades de análisis descriptivo sobre los datos históricos (lo que sucedió) de los especialistas en este negocio. No obstante, aún se carece de capacidades de análisis diagnóstico (por qué sucedió), predictivo (qué sucederá) y prescriptivo (qué debería hacer) que permitan mejorar el desempeño del proceso de negocio, debido principalmente a aspectos relacionados con la disponibilidad y la calidad de la información que se produce en todas sus etapas.

Este trabajo aborda de manera general la problemática del manejo de equipajes en la industria aeronáutica, presenta una revisión de los trabajos previos relacionados con la misma y plantea el diseño e implementación de una herramienta de analítica visual como alternativa de solución, desde el punto de vista del manejo y análisis de información. La primera parte presenta el marco teórico y estado del arte. La segunda parte describe el planteamiento del problema y los objetivos propuestos. En la tercera parte se presenta la metodología empleada. En la cuarta parte se describen los resultados obtenidos y por último, en el capítulo quinto, se presentan las conclusiones y recomendaciones del trabajo realizado.

Capítulo 1. Marco teórico

En esta sección se introducen los conceptos más relevantes del proceso de gestión de equipajes, los tipos de irregularidades, la relación con el mercado de transporte aéreo de pasajeros y las regulaciones asociadas.

1.1. Contexto de la Industria

1.1.1 Cifras del mercado global

De acuerdo con el reporte anual de equipajes de SITA, uno de los líderes del mercado de telecomunicaciones y soluciones de tecnología para la industria aeronáutica (SITA, 2019a), desde el año 2007 el mercado de transporte aéreo de pasajeros ha mantenido un incremento constante, no así las irregularidades asociadas al equipaje (Figura 1).



Figura 1. Tendencia de pasajeros transportados e incidentes de equipajes 2007-2017.

Esto se debe principalmente a la implementación de sistemas BRS (*Baggage Reconciliation System*) para el registro y seguimiento de equipajes en los aeropuertos. Sin embargo, reducir la cifra de irregularidades continúa siendo un objetivo primordial en la industria aeronáutica global, considerando que para el 2017 los costos asociados ascendieron aproximadamente a 2.3 billones de dólares (SITA, 2018), representados, entre otros, en pagos por compensación a los pasajeros.

A pesar de la implementación de sistemas BRS, aún existen diferentes puntos en el proceso de gestión de equipajes donde se presentan situaciones que dan lugar la generación de irregularidades, especialmente por la existencia de tareas manuales y por la participación de diferentes compañías durante la operación del proceso, como el personal de concesión de los aeropuertos, el personal de las aerolíneas, los servicios de operaciones terrestres tercerizados en otras organizaciones, entre otros.



Figura 2. Causas de equipaje retrasado e incidentes presentados en 2017.

Las irregularidades reportadas por los pasajeros se clasifican en cuatro tipos: demora, pérdida, daño y saqueo. Las más frecuentes son las demoras con un 78% de los casos, le siguen los daños y saqueos con un 17% y las pérdidas con un 5%. Por otro lado, la identificación y asignación de la(s) causa(s) atribuible(s) a cada reclamo es el resultado de procedimientos de investigación propios de aeropuertos y aerolíneas. Dentro de las causas más frecuentes se encuentran los errores en transferencias y conexiones (47%), otros

errores menos frecuentes están relacionados con el etiquetado del equipaje, la carga hacia la aeronave, errores en la llegada, condiciones climáticas, restricciones por peso y espacio, entre otros (Figura 2).

Estas causas se encuentran catalogadas en un listado conocido como Códigos RL (*Reason for Loss Codes*), el cual maneja una codificación principal para los tipos de causa y un código secundario que representa la causa específica (Farey, 2014), como se ilustra en la Figura 3.

REASON FOR LOST CODES		
	<i>P</i> = PRIMARY CODE	<i>S</i> = SECONDARY CODE
P	10 STATION ERRORS - TAGGING	
S	11 INCORRECT OR NO ENTRIES ON TAG	
	12 NOT CHECKED TO FINAL DESTINATION	
	14 BAGGAGE NOT LABELED	
	15 WRONG BAG LABELED- e.g. TAG SWITCH	
	17 OLD TAG NOT REMOVED	
	18 BAGGAGE NOT RECONCILABLE / WRONG DATA IN CHECK-IN SYSTEM	
P	20 STATION ERRORS - SHORTSHIPPED	
S	21 BAGGAGE LEFT AT STATION IF ORIGIN / CORRECTLY LABELED	
	22 BAGGAGE RECEIVED TOO LATE FROM CHECK-IN	
	23 STANDBY BAGGAGE LEFT BEHIND	
	24 BAG CHECKED-IN AIRSIDE AND FAILED TO LOAD	
	25 GATE CHECKD-IN BAGGAGE LEFT BEHIND	
	26 OFFLOADED DUE SPACE / WEIGHT RESTRICTION	

Figura 3. Ejemplo de Códigos RL.

1.1.2 Proceso general: *The baggage journey*

El proceso de gestión de equipajes involucra el transporte tanto para los vuelos de salida (desde la recepción del equipaje en el *counter* hasta la carga hacia la aeronave) como para los vuelos de llegada (desde la descarga del equipaje hasta la entrega a los pasajeros). Las conexiones o transferencias entre aerolíneas son un subconjunto de estas etapas: se recibe el equipaje de una aeronave (descarga) y se conduce hacia la otra (carga), para luego ser entregado al pasajero en el aeropuerto destino. Este proceso se ilustra en la Figura 4 (SGR-Comunicaciones, 2018).

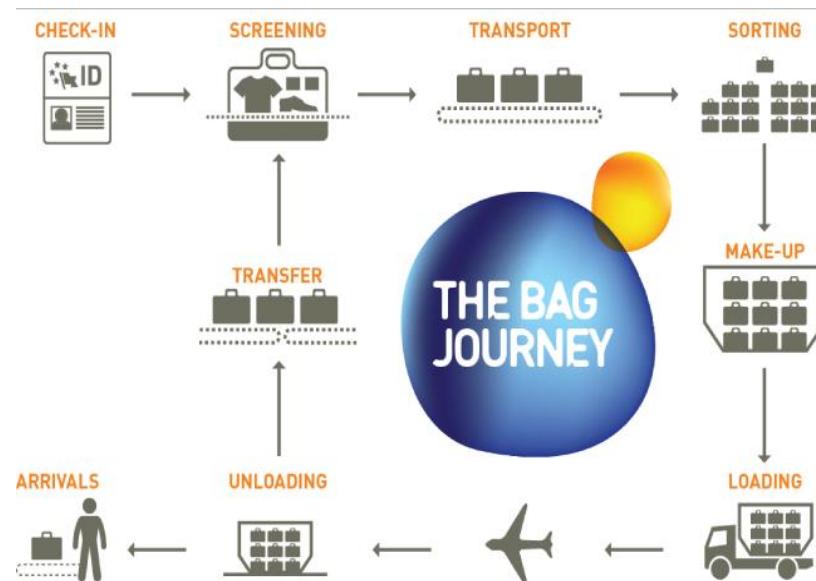


Figura 4. The baggage journey.

Las etapas que recorre el equipaje en este proceso se describen a continuación: en primer lugar el equipaje es entregado por el pasajero en el punto de check-in de la aerolínea, donde recibe un identificador único (*bag tag*) que está asociado al tiquete de vuelo del pasajero; luego es revisado en puntos de seguridad con scanners de rayos X para identificar elementos no permitidos de acuerdo a la normatividad; posteriormente es conducido por el sistema de manejo de equipajes BHS (*Baggage Handling System*) a un punto (*sorting area*) en el cual el sistema separa automáticamente el equipaje de acuerdo a la aerolínea en la que será transportado, utilizando para ello lectores del código asignado en el *bag tag*. A continuación el equipaje es organizado manualmente en contenedores o directamente en los vehículos según el vuelo al que pertenecen (*make-up*), y desde allí es transportado para ser cargado en la aeronave correspondiente (*loading*). Cuando el equipaje pertenece a un vuelo con conexión, se descarga el contenedor (*unloading*) y se lleva a una zona de transferencia desde donde es enrutado hacia el vuelo correspondiente (*transfer*), en caso contrario, el equipaje es descargado y conducido a la zona donde los pasajeros finalmente lo reclaman (*arrivals*). El proceso también presenta otros puntos de gestión adicional de acuerdo con la normativa de seguridad propia de cada aerolínea y aeropuerto.

1.1.3 Puntos de registro obligatorios

La Asociación Internacional de Transporte Aéreo (IATA, por sus siglas en inglés), por medio de la resolución 753, establece que las aerolíneas deben estar en capacidad de:

- Demostrar la entrega de equipaje, cuando trasladan la custodia
- Demostrar la adquisición del equipaje, cuando reciben la custodia
- Proveer un inventario de equipajes a la salida de un vuelo

Para el cumplimiento de dichas obligaciones, se establecen cuatro puntos mínimos de registro y seguimiento de equipaje: aceptación, carga, transferencia y llegada (Figura 5) (IATA, 2019a).



Figura 5. Puntos de registro obligatorios en *The Baggage Journey*.

La normativa no reglamenta el uso de tecnologías para esta labor, el registro puede realizarse manual o automáticamente, no obstante se enumeran algunas tecnologías de uso potencial como lectores láser, RFID, reconocimiento óptico de caracteres (OCR), bluetooth, NFC, Wi-Fi, GPS, entre otros. Adicionalmente, se identifican posibles ubicaciones para realizar el seguimiento, especialmente donde se presenta cambio en la custodia del equipaje, de acuerdo con cada punto obligatorio de registro.

A través de la implementación de estos puntos de registro y seguimiento, la IATA pretende generar beneficios en términos de un mejor análisis de causa-raíz de irregularidades, reducción de fraudes, calificación favorable de los clientes, medición del rendimiento del proceso y mejor experiencia del pasajero (IATA, 2017).

1.1.4 Indicadores del proceso

El principal indicador de desempeño del proceso se conoce como MB (*Mishandled Baggage*), el cual mide el número de irregularidades presentadas por cada mil pasajeros transportados. Se calcula de la siguiente manera: (número de irregularidades/pasajeros) *1000. Un referente del mercado se encuentra en (Proceedings, 2018), el incluye, entre otros, el reporte de irregularidades de equipaje para aerolíneas de Estados Unidos, organizado por el indicador MB (**¡Error! No se encuentra el origen de la referencia.**).

Tabla 1. Reporte MB para aerolíneas americanas a octubre de 2018.

RANK	AIRLINE	OCTOBER 2018		
		TOTAL BAGGAGE REPORTS	ENPLANED PASSENGERS	REPORTS PER 1000 PASSENGERS
1	DELTA AIR LINES	13,871	10,630,284	1.30
2	SPIRIT AIRLINES	2,921	2,099,421	1.39
3	JETBLUE AIRWAYS	3,894	2,686,996	1.45
4	UNITED AIRLINES	15,595	7,966,365	1.96
5	FRONTIER AIRLINES	3,687	1,686,069	2.19
6	ALASKA AIRLINES**	6,741	2,756,891	2.45
7	SOUTHWEST AIRLINES	34,748	13,908,773	2.50
8	HAWAIIAN AIRLINES	2,121	847,341	2.50
9	AMERICAN AIRLINES	32,383	9,698,949	3.34
10	SKYWEST AIRLINES	8,539	2,389,666	3.57
11	EXPRESSJET AIRLINES	2,346	590,911	3.97
12	ENVOY AIR	5,341	1,065,170	5.01
	TOTALS	132,187	56,326,836	2.35

Considerando el costo que estas irregularidades representan para la industria y la percepción de calidad del servicio por parte de los pasajeros, otros indicadores propios de cada compañía pueden estar asociados a los costos generados a causa de los pagos por compensación y el NPS (*Net Promoter Score*), respectivamente.

1.2. Alternativas para el mejoramiento del proceso de equipajes

En esta sección se presenta el estado del arte de la temática de manejo de equipajes, organizado según la perspectiva de los estudios previos que han abordado el tema y los sistemas de información existentes.

1.2.1. Modelamiento, optimización y simulación

Diversos estudios han abordado el tema de manejo de equipajes buscando optimizar el proceso de manera parcial o total. En (Alsyouf, Humaid, & Al Kamali, 2014) se aplicaron metodologías como six sigma y análisis cuantitativo y cualitativo sobre los datos de cada etapa del proceso, con hallazgos relevantes sobre las causas que generan irregularidades de equipajes, siendo las más destacadas: 1) la ausencia de entrenamiento adecuado del personal que manipula los equipajes durante la transferencia en vuelos con conexión, 2) las largas jornadas de trabajo del personal del proveedor de servicio de equipajes, 3) las fallas en el sistema de bandas transportadoras y 4) las caídas de las maletas del sistema de bandas por diversos factores como la inclinación, velocidad y manipulación inadecuada. También se han propuesto alternativas de solución a cada uno de estos problemas, que incluyen el incremento en los niveles de entrenamiento y evaluación de desempeño al personal, la aplicación de algoritmos de programación lineal para reducir las jornadas de trabajo de 12 a 8 horas, la aplicación de metodologías de análisis de efecto y modo de fallos (FMEA, por sus siglas en inglés) para introducir tecnologías de mantenimiento predictivo a través del monitoreo de las condiciones del sistema BHS, y la propuesta de ajuste en los ángulos de inclinación del sistema de bandas para minimizar las caídas de las maletas.

(Frey, 2014) estudia el manejo del equipaje en los aeropuertos desde una perspectiva de investigación de operaciones (OR). La planificación y el control del flujo de equipaje en los puntos de entrada y puntos de salida de los sistemas BHS (Figura 6) son planeadas principalmente por trabajadores experimentados, llamados despachadores. Su tarea es planificar el manejo del equipaje de manera que los recursos dados se utilicen de manera eficiente. Sin embargo, debido a la gran cantidad de información, por ejemplo la cantidad de equipaje que se transferirá a un vuelo o el horario de vuelo, la tarea de los despachadores es un desafío. Para mejorar la planificación del manejo de equipaje en los aeropuertos, este estudio aplica técnicas de OR avanzadas para controlar el flujo de equipaje en los extremos del sistema BHS para el proceso de manejo de equipaje de entrada y salida. Se plantea un primer subproceso que describe la transferencia de equipaje de los pasajeros o vuelos entrantes a los vuelos que salen, mientras que el segundo subproceso planea la transferencia de equipaje de los vuelos entrantes a los pasajeros. Al utilizar algoritmos para la planificación, aplicados sobre datos del mundo real, se evidencia un mejor control del manejo de equipaje con una planificación más robusta y eficiente, haciendo más predecibles los errores o la mala planificación.

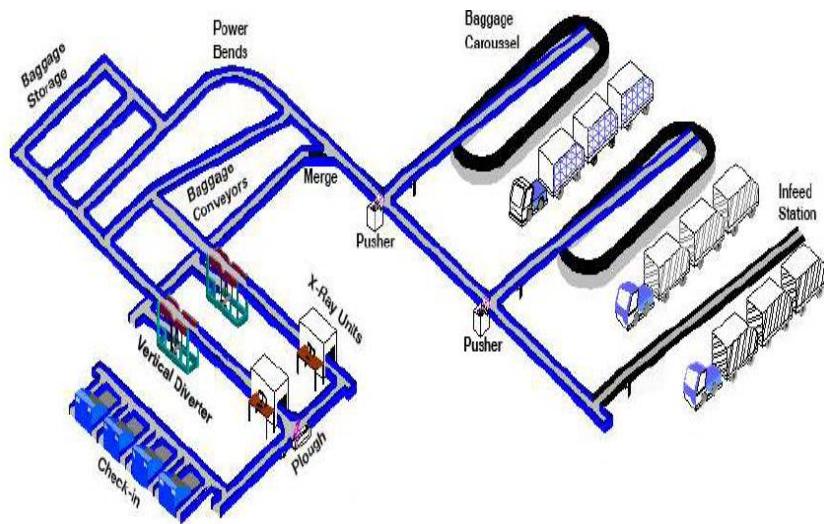


Figura 6. Un sistema BHS y sus componentes.

Otros estudios plantean alternativas en términos de modificaciones a las etapas del proceso desde el punto de vista de la interacción con el pasajero. Se tiene como referente un caso de estudio realizado en el aeropuerto Schiphol en Ámsterdam, donde se propone un modelo basado en la implementación del servicio de check-in fuera del aeropuerto (van Zundert, 2010), buscando descongestionar los puntos de recepción de equipaje, a través de la oferta de cinco servicios diferentes: 1) servicio de recogida de equipaje: el equipaje se recoge en una dirección especificada por el pasajero la noche anterior a la salida; 2) servicio de entrega local: el equipaje puede facturarse en un punto de entrega local, el día anterior a la salida; 3) servicio en la estación de tren: el equipaje se puede facturar en las siete estaciones de tren principales de los Países Bajos hasta tres o cuatro horas antes de la salida; 4) servicio de estacionamiento de Schiphol: el equipaje se puede registrar en los estacionamientos de larga estancia de Schiphol hasta una hora y media antes de la salida; y 5) servicio en la estación Schiphol: el equipaje puede registrarse en la estación de tren de Schiphol hasta una hora y media antes de la salida.

De otro lado, diversos trabajos han abordado la problemática en la etapa final del proceso: la recepción del equipaje de los vuelos entrantes y su correspondiente entrega a los pasajeros (*inbound baggage*). El trabajo de (Frey, Kiermaier, & Kolisch, 2017) establece que el balance en el uso de las bandas transportadoras y la reducción de los tiempos de espera de los pasajeros en la zona de entrega de equipaje son dos objetivos fundamentales de este subproceso. El aporte realizado describe de manera detallada el subproceso de manejo de equipaje entrante (Figura 7), realiza un modelamiento matemático del problema,

establece su complejidad y aplica metodologías como GRASP (*greedy randomized adaptive search procedure*), GFLS (*guided fast local search*) y *path-relinking*; el resultado de la aplicación de este modelo al aeropuerto de Munich permitió reducir los picos de congestión de equipaje en las bandas en un 38% (ocupación máxima de la banda en un 80%) y reducir los tiempos de espera de los pasajeros en un 11% (tiempo promedio de espera en 5.9 minutos).

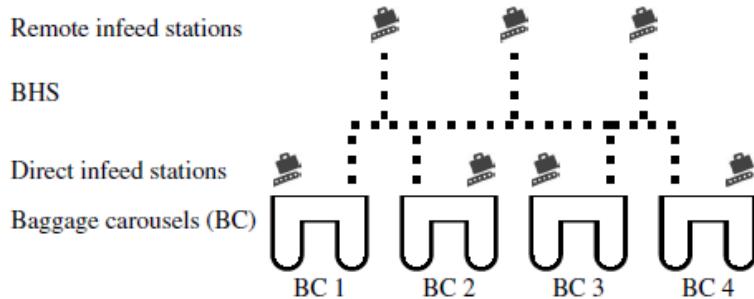


Figura 7. Infraestructura para el subproceso de equipaje entrante.

Adicionalmente, se han realizado simulaciones del subproceso de equipaje entrante con software comercial especializado para esta labor. En (Malandri, Briccoli, Mantecchini, & Paganelli, 2018) se descompone el subproceso en cada uno de sus pasos, se identifican los tiempos invertidos en cada paso (Figura 8) y se realiza la medición utilizando como caso de estudio el aeropuerto de Bologna, en Italia, por un periodo de dos semanas.

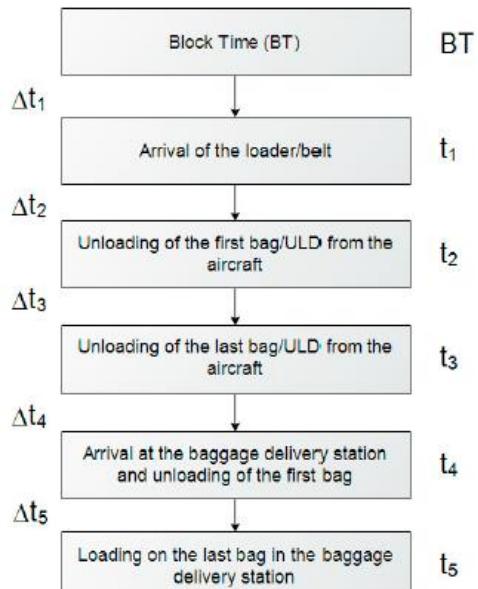


Figura 8. Puntos y tiempos en el subproceso de manejo de equipaje entrante.

A través del uso del software comercial Arena, se realiza la simulación del subprocesso de manejo de equipaje entrante y se obtienen resultados satisfactorios entre la información simulada y la real, así como factores de interés que pueden causar discrepancias en este tipo de estudios, como las temporadas pico en los aeropuertos, los tipos de vuelo asignados a las bandas de transporte (en este caso los que provienen de países dentro del acuerdo Schengen y los que no) y las estaciones climáticas.

El uso de tecnologías más sofisticadas como los vehículos con código de destino (DCV-*destination coded vehicles*, vehículos de alta velocidad propulsados por motores de inducción lineal, que transportan las piezas de equipaje desde las estaciones de carga a las estaciones de descarga), impone retos importantes en las tareas de asignación y programación (Figura 9) (Alstef, 2019b). Desde la perspectiva de optimización del uso de estos vehículos, se han realizado trabajos que abordan el problema de control de los DCVs en los sistemas de manejo de equipajes.



Figura 9. Vehículos con código de destino (DCV).

El trabajo de (Zeinaly, De Schutter, & Hellendoorn, 2013) identifica tres problemas de control fundamental: programación y enrutamiento, balanceo de carga y manejo de vehículos vacíos. Para esto se realiza el modelamiento como problemas de programación lineal (LP) y no lineal (NLP), buscando que las piezas de equipaje lleguen a la estación de descarga dentro de la ventana de tiempo preespecificada, y que se minimice el costo de operación del sistema (consumo de energía y almacenamiento del equipaje). Finalmente, se demuestra por medio de simulaciones que el modelo propuesto alcanza un desempeño favorable con un esfuerzo computacional relativamente bajo.

1.2.2. Análisis de información espacio-temporal y en movimiento

Este apartado trata los estudios realizados en términos del análisis de datos. Considerando la temática del manejo de equipajes, es completamente posible analizar una pieza de equipaje como un objeto discreto, que a través del movimiento inducido realiza un recorrido y ocupa diversas ubicaciones en el espacio en diferentes momentos del tiempo. Los datos que pueden ser analizados utilizando esta aproximación enfrentan diversos desafíos en su representación, almacenamiento, procesamiento y visualización. Para el abordaje de la problemática de irregularidades de equipajes desde una perspectiva de análisis de información, es relevante considerar este tipo de trabajos.

En el estudio de (Andrienko, Andrienko, & Gatalsky, 2003) se realiza una revisión de las técnicas basadas en visualización para el análisis exploratorio de datos espaciotemporales, y a partir del resultado propone una tipología de tareas de análisis (Figura 10).

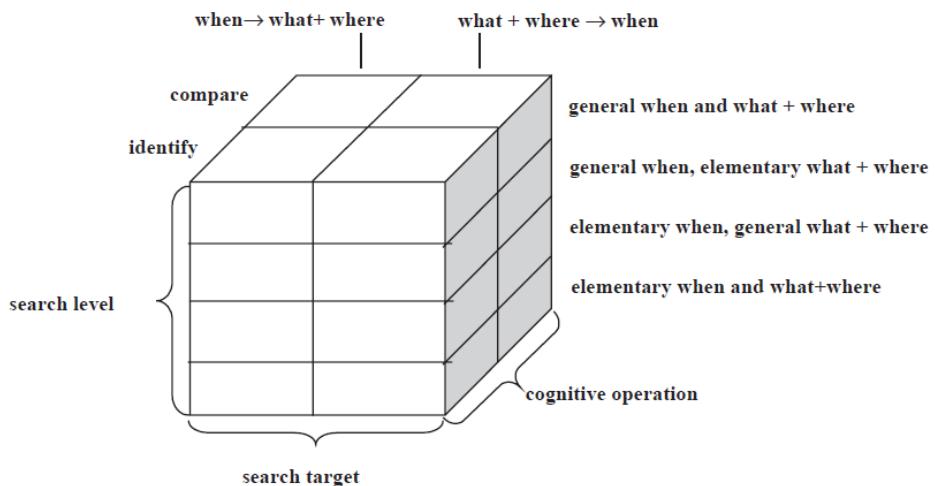


Figura 10. Tipología de tareas para análisis de datos espaciotemporales.

Esta tipología está conformada por tres dimensiones: el objetivo de búsqueda, los niveles de búsqueda y las operaciones cognitivas. Adicionalmente, los elementos de datos son representados por medio de tres atributos: qué (objeto), cuándo y dónde (*what, when, where*). Los objetivos de búsqueda se plantean en términos de estos tres atributos, así una búsqueda podría preguntar por la ubicación de un objeto en un instante específico de tiempo, o el momento del tiempo en el que un objeto ocupaba un lugar específico. Las operaciones cognitivas se resumen en identificar (para elementos puntuales) y comparar (para conjuntos de elementos). Así mismo, los niveles de búsqueda se clasifican en

elementales (para elementos individuales) y generales (para conjuntos). De esta manera se origina un amplio conjunto de tareas de análisis derivadas de la combinación de estas dimensiones, conformando un insumo importante para el diseño de herramientas de visualización y análisis exploratorio de datos.

Otro referente importante en el análisis de datos espacio-temporales se encuentra en (Chen et al., 2017), el cual describe las técnicas utilizadas para el diseño e implementación de VAUD (*Visual Analyzer for Urban Data*) una herramienta de analítica visual para la exploración de datos urbanos.

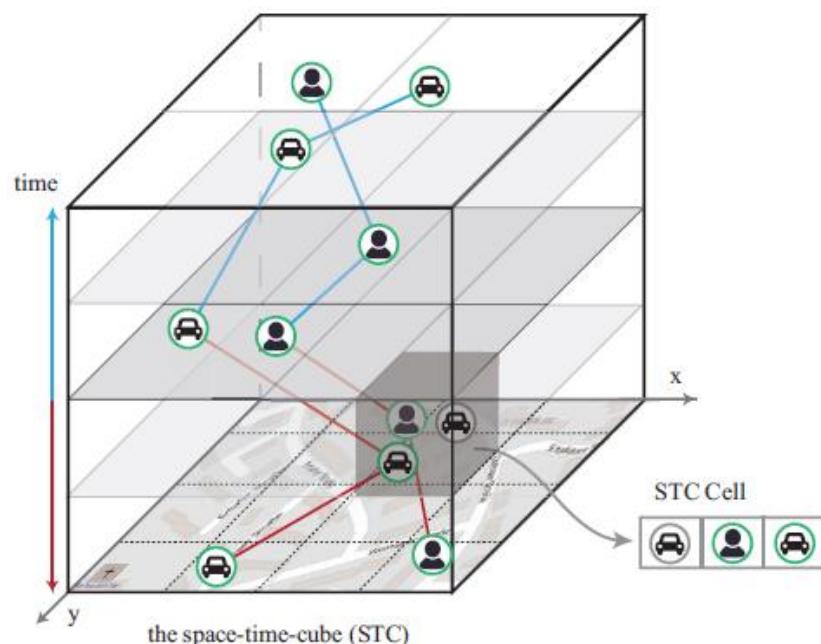


Figura 11. Cubo de espacio-tiempo.

En este estudio se utiliza el cubo de espacio-tiempo para representar los datos, siendo X y las coordenadas de espacio, y la altura la representación del tiempo dividida en porciones (rebanadas) como se muestra en la Figura 11. También emplea los atributos de espacio (*where*), tiempo (*when*), objeto (*what*) y agrega un atributo de identificación (*which*). Con estos elementos se diseña un modelo de consultas (*queries*) basado en modismos de visualización y se pone a prueba la herramienta con datos urbanos reales, obteniendo resultados satisfactorios por parte de los usuarios.

Un referente más cercano al problema de manejo de equipajes se encuentra en (Al-Serafi & Elragal, 2014). Allí se aborda el problema desde una perspectiva de minería de datos para el análisis de trayectorias de equipajes dentro de un sistema BHS (Figura 12).

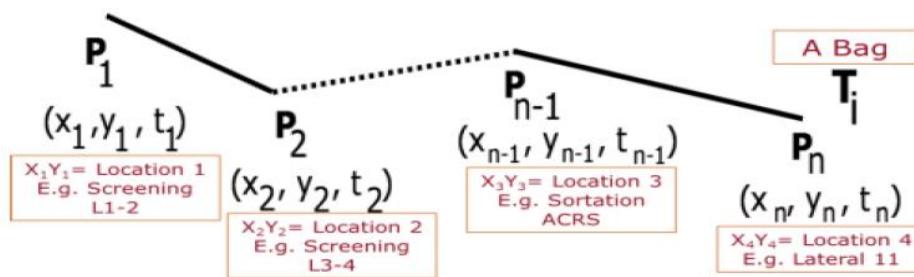


Figura 12. Trayectoria de una pieza de equipaje dentro del sistema BHS.

Este trabajo propone una arquitectura en capas para la implementación de una herramienta que soporte este tipo de análisis (Figura 13), y realiza énfasis en la capa de visualización para el análisis de trayectorias, estableciendo los objetivos principales de análisis para el caso de estudio de manejo de equipajes y describiendo las técnicas de análisis visual empleadas para responder a cada objetivo.

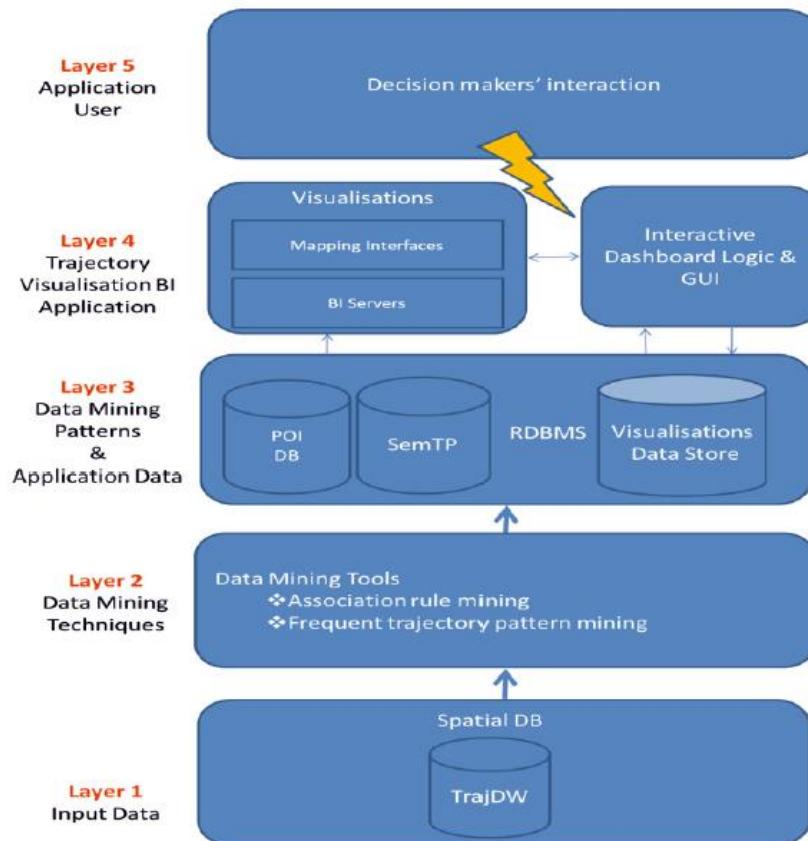


Figura 13. Marco de trabajo para el análisis de trayectorias.

Esta aproximación resulta en una herramienta que permite abordar problemas comunes en el manejo de equipajes, como la identificación de los puntos del sistema BHS de uso más

común, los patrones de trayectorias más frecuentes desde y hacia una ubicación específica y la identificación de congestión en el sistema basado en las frecuencias de visitas en momentos específicos del tiempo. Este tipo de análisis se realizan de manera visual por medio de la interfaz de usuario de la herramienta (Figura 14).

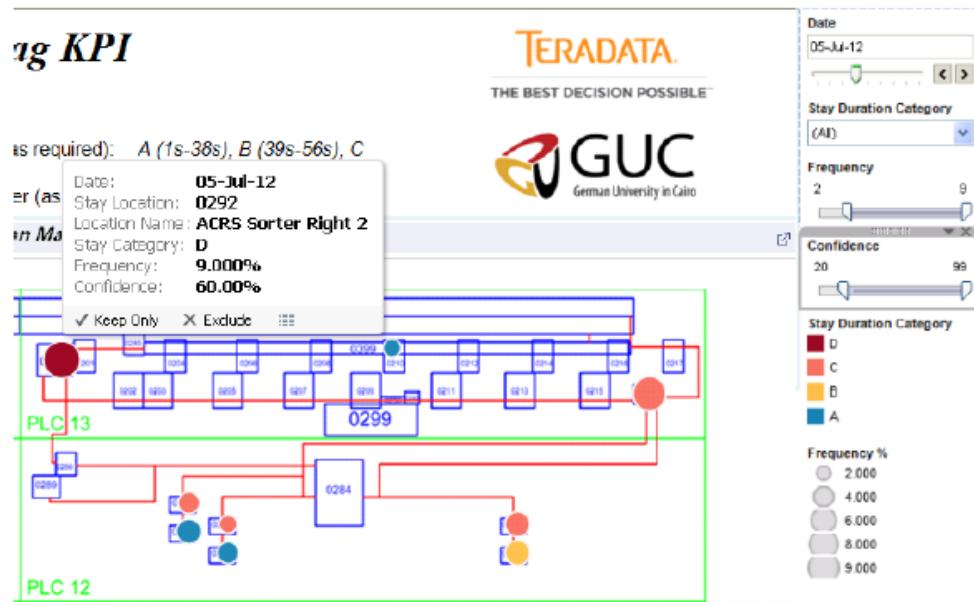


Figura 14. Visualización de puntos de congestión.

1.2.3. Sistemas de manejo de equipajes

Como parte de la revisión es fundamental identificar los principales sistemas de información de manejo de equipajes que son utilizados en la actualidad, los cuales se describen a continuación:

A. Sistemas de manejo de equipaje (*Baggage Handling System - BHS*)

Sistema conformado por una red de cintas transportadoras que mediante un sistema automatizado trasladan el equipaje dentro del sistema aeroportuario (Wikipedia, 2018). Estos sistemas tienen un conjunto de elementos estandarizados en general de acuerdo con el proceso de manejo de equipajes, sin embargo su disposición es altamente dependiente de la infraestructura de cada aeropuerto (Figura 15) (ALSTEF, 2019a).

B. Sistema de conciliación de equipaje (*Baggage Reconciliation System - BRS*)

Es un sistema utilizado para registrar y rastrear la ubicación del equipaje en todo su recorrido, desde el registro en el punto de *check-in* hacia la aeronave para los vuelos

de salida, y en la descarga y transferencia de equipajes en vuelos de llegada o en conexión. Está conformado principalmente por etiquetas que se asignan a las piezas de equipaje (*bag tag*), dispositivos de lectura manuales o automáticos en puntos específicos del aeropuerto, y un sistema de información para la gestión y consulta de dichos datos. Dos de los proveedores más reconocidos en el mercado son (Amadeus, 2019c) y (SITA, 2019b).

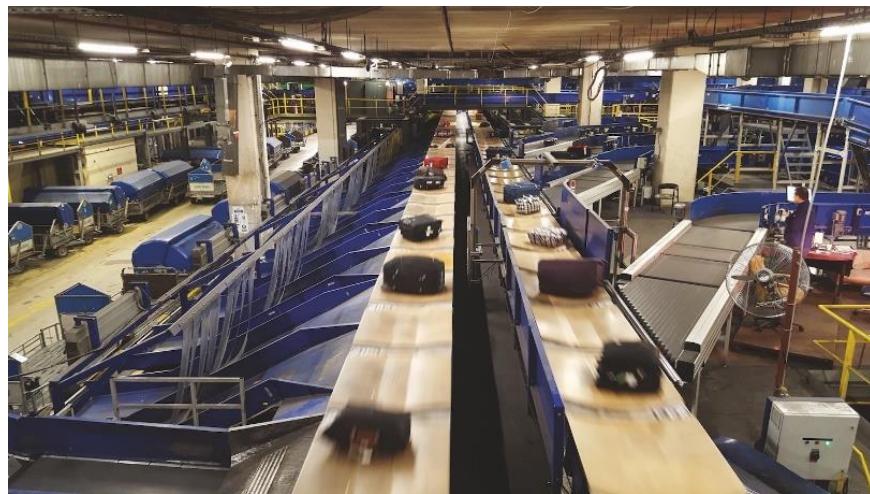


Figura 15. Ejemplo de sistema BHS.

C. Sistemas de rastreo y búsqueda de equipaje

Actualmente existe un servicio proporcionado por el sistema WorldTracer (SITA, 2019c), para el rastreo de equipaje perdido o demorado. Ha estado en operación por varios años y es utilizado por todas las aerolíneas que pertenecen a la IATA (Wikipedia, 2019b).

1.2.4. Herramientas para diseño de soluciones de analítica e inteligencia de negocios

Considerando el contexto de los sistemas de manejo de equipajes existentes y la necesidad de alternativas de solución desde el punto de vista de análisis de información, se hace necesaria la revisión de las herramientas disponibles para construir soluciones que permitan mejorar las capacidades de análisis de esta industria. El enfoque de esta revisión se hará en las herramientas que brindan capacidades de análisis visual interactivo, no obstante existen otras herramientas que abarcan capacidades de almacenamiento,

procesamiento, manejo de grandes volúmenes de datos, virtualización, entre otras, que no son parte del alcance de la revisión.

De acuerdo con el cuadrante de Gartner para plataformas de inteligencia de negocios y analítica (Howson, Richardson, Sallam, & Kronz, 2019) las tres mejores propuestas provienen de Microsoft, Tableau y Qlik (Figura 16).



Figura 16. Cuadrante de Gartner para plataformas de BI y analítica.

Power BI es un servicio de análisis de negocios que brinda información para permitir decisiones rápidas e informadas (Microsoft, 2019b). Dentro de sus fortalezas se encuentran la mejor relación de costo-beneficio entre sus competidores, la facilidad de uso para análisis complejos y la visión de producto, que contempla dentro de su *roadmap* la adición de capacidades analíticas adicionales a través de servicios de la nube de Azure (Howson et al., 2019), como *Azure Machine Learning* y *cognitive services*.

Tableau es una compañía de software de visualización interactiva de datos (Tableau, 2019), su fortaleza radica en ofrecer una experiencia intuitiva que permite a los usuarios acceder,

preparar, analizar y presentar hallazgos en sus datos, sin conocimientos técnicos o de programación.

Qlik es una compañía reconocida por sus productos QlikView y Qlik Sense, cuyo propósito es la visualización de datos y la inteligencia de negocios; QlikView es para analítica guiada y Qlik Sense es para visualización en modo autoservicio (Qlik, 2019). Dentro de sus fortalezas está la extensibilidad, ya que permite construir aplicaciones personalizadas a través de un amplio conjunto de APIs.

Estas tres alternativas coinciden en brindar capacidades de análisis en modo autoservicio, con curvas de aprendizaje relativamente cortas para cualquier usuario de negocio. Contienen modismos de visualización de uso tradicional (*table, pie charts, line charts, bar charts, stacked bar charts, scatter plots, bubble charts, area, stacked area, star plot*) y también otros más robustos (*funnel, maps, tree maps, heat maps, stream graphs, histogram, box plot, Gantt view*). Algunas de ellas permiten extensiones y configuraciones para incluir modismos de uso en análisis más complejos (*parallel coordinates, networks, trees, sunburst*) y también la adición de objetos visuales y herramientas de tratamiento de datos generadas en otras tecnologías, como scripts de Python o R. Adicionalmente, disponen de conectores hacia una gran variedad de fuentes de datos.

Los estudios previos sobre optimización del proceso de manejo de equipajes permiten obtener un entendimiento suficiente de su comportamiento. Los diferentes modelos planteados contribuyen a la identificación, medición y mejora de puntos críticos del proceso. Sin embargo, el entendimiento de la problemática global de las irregularidades de equipajes, la identificación de sus causas y las posibles acciones de mejora se ve limitado a las etapas o subprocessos analizados (optimización en *check-in*, recepción de *inbound baggage*).

Respecto a los sistemas de manejo de equipajes existentes, presentan un enfoque netamente transaccional, centrado en el soporte a la operación de aeropuertos y aerolíneas, con cobertura a etapas específicas del proceso de manejo de equipajes. No todos los aeropuertos disponen de estos sistemas, especialmente los BRS, debido a su costo. Cada aerolínea complementa las funcionalidades y la información para analizar con sistemas adicionales de soporte al proceso, como los sistemas de gestión del relacionamiento con los clientes (CRM) y los sistemas de la operación de vuelos.

Desde el punto de vista de análisis de información, estos sistemas manejan reportes en mayor o menor grado, sin embargo, no toda la información es accesible, oportuna y sin costo. Adicionalmente, considerando que la cobertura se da sobre etapas puntuales del proceso, no se logra a través de ellos reunir la totalidad de información que permita analizar el proceso de inicio a fin. Por esta razón es común encontrar implementaciones de reportes personalizados en distintas áreas y con diferentes herramientas, o diseño de soluciones de inteligencia de negocios tradicional, bodegas de datos, *datamarts* y reportes con visualizaciones personalizadas, en compañías con mayores necesidades y también con mayor presupuesto. No obstante, la mayoría de estas implementaciones cubren las capacidades de análisis descriptivo, es decir, los hechos ocurridos, a partir de los datos históricos con las limitaciones ya mencionadas. Las capacidades de análisis diagnóstico (por qué sucedieron los hechos), predictivo (qué sucederá) y prescriptivo (qué acciones se deben tomar) son ausentes bajo estas condiciones.

Las herramientas para el diseño de soluciones analíticas y de inteligencia de negocios presentan fortalezas importantes en la obtención de datos desde múltiples fuentes y en el uso de elementos de análisis visual, enfocado a un modelo de autoservicio de fácil uso por parte de los usuarios de negocio. No obstante, es necesario profundizar en la cobertura que los diferentes modismos de visualización ofrecidos pueden tener en el análisis de datos en movimiento y con características espaciotemporales, especialmente cuando son enmarcados en un dominio específico de negocio con necesidades particulares, como aquellos asociados al entendimiento y reducción de las irregularidades de equipajes.

Con base en esta revisión, se identifica que una alternativa de solución debería contar con las siguientes características:

- *Cobertura integral del proceso de manejo de equipajes.* En términos de la completitud de las etapas del proceso y los puntos que hacen parte del recorrido de las piezas de equipaje. También en considerar los datos de cada una de estas etapas del proceso, y de procesos estrechamente relacionados asociados como la operación de vuelos.
- *Soporte a capacidades analíticas avanzadas.* Incluir no solo capacidades de análisis descriptivo tradicional, sino también capacidades de análisis diagnóstico y predictivo.

- *Enfoque en el usuario experto del negocio.* Incorporar capacidades de analítica visual que permitan aumentar los tipos de análisis del usuario a través de interfaces y modismos de visualización de fácil uso, adaptados al dominio del problema.

Capítulo 2. Planteamiento del Problema y Objetivos

2.1. Planteamiento del problema

La industria de transporte aéreo de pasajeros representa un mercado creciente con una operación influenciada por diversos factores que dificultan la consecución de sus objetivos (clima, conexiones, seguridad, infraestructura aeroportuaria), con una problemática particular en el manejo de equipajes que suma 2.3 billones en costos asociados a 22.7 millones de casos presentados para el año 2017, relacionados con demoras, daños, pérdidas o saqueos. Adicionalmente, no se cuenta con sistemas y soluciones que aborden el proceso de manejo de equipajes de manera integral y las capacidades analíticas actuales están limitadas al análisis descriptivo.

Con base en dicha problemática, se plantea la siguiente pregunta: Cómo mejorar el análisis de irregularidades en el manejo de equipajes a través del uso de una herramienta de analítica visual?

2.2. Objetivo General

Diseñar una herramienta de analítica visual que permita apoyar el análisis de irregularidades en el manejo de equipajes.

2.3. Objetivos Específicos

1. Identificar las problemáticas y necesidades más comunes en el proceso de gestión de equipajes.

2. Realizar el modelamiento de los datos y determinar los tipos de análisis y técnicas acordes al problema identificado.
3. Diseñar e implementar una herramienta de analítica visual que permita realizar análisis descriptivos y diagnósticos del manejo de equipajes.

Capítulo 3. Metodología

Considerando los objetivos propuestos, se plantean algunos marcos de referencia para el abordaje metodológico en términos de las capacidades analíticas, las etapas de un proyecto de ciencia de datos y los marcos de trabajo para el desarrollo de soluciones de analítica visual.

3.1. Capacidades analíticas

En primer lugar, se pretende abordar las capacidades de analítica descriptiva y diagnóstica, como se ilustra en la Figura 17 (Sapp C., 2018).

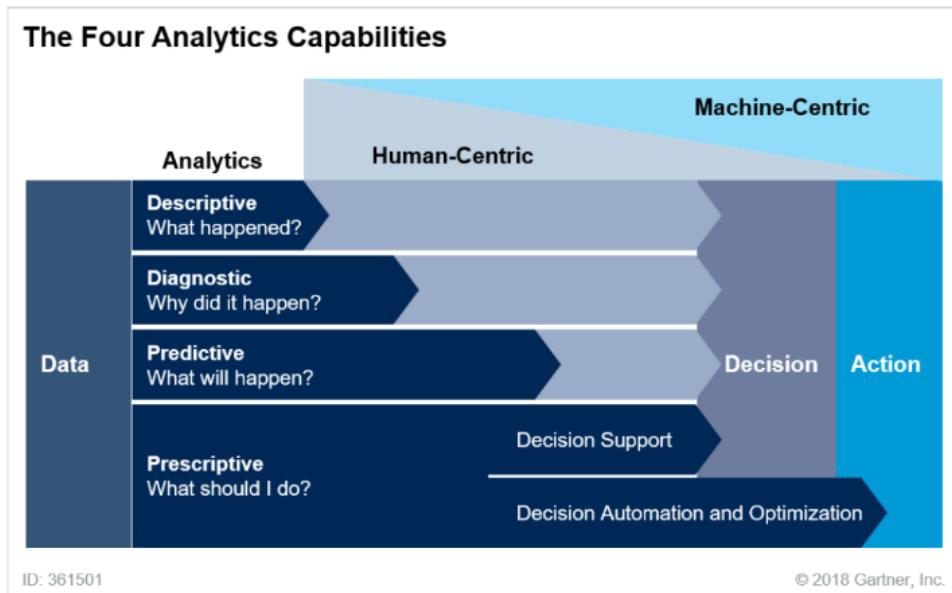


Figura 17. Las 4 capacidades analíticas.

La analítica descriptiva se centra en la descripción de los hechos ocurridos que son representados por medio de los datos (qué, cuándo, cuántos y cómo sucedieron). La analítica diagnóstica se enfoca en la explicación de las razones por las cuales sucedieron

esos hechos, incorporando comúnmente fuentes de datos adicionales y relacionados al dominio de negocio en estudio. La analítica predictiva tiene como fin realizar predicciones de los hechos que sucederán a partir del análisis de los datos históricos, por medio del uso de técnicas que utilizan parte de estos datos como conjunto de entrenamiento y otra parte como datos para probar los modelos. Por último, la analítica prescriptiva pretende sugerir acciones a partir de los datos que apoyan la toma de decisiones, ya sea de manera manual o automatizada.

3.2. Metodología ASUM-DM

De otro lado, considerando el enfoque general del trabajo en análisis de datos, se utilizó como guía los pasos de la metodología ASUM-DM (Rollins, 2015) (Figura 18). Basada a su vez en la metodología CRISP-DM y enfocada en proyectos de ciencia de datos, está conformada por 10 etapas que representan un proceso iterativo completo que abarca desde la concepción de la solución hasta su despliegue y retroalimentación.

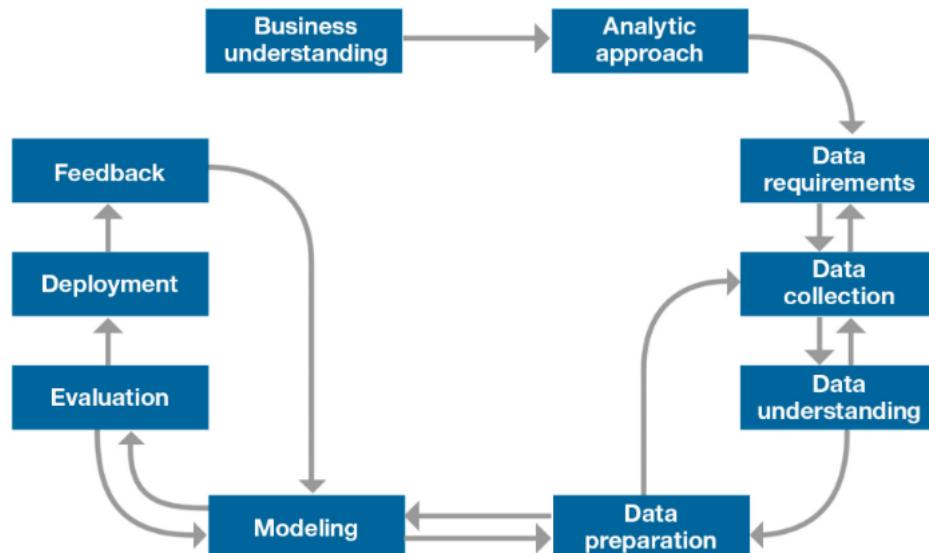


Figura 18. Metodología ASUM-DM.

En la primera etapa (**Entendimiento del negocio**) se define el problema, los objetivos y requerimientos de la solución desde una perspectiva de negocio.

En la etapa de **Aproximación analítica** se expresa el problema en el contexto de técnicas (usualmente estadísticas, matemáticas o de aprendizaje automático) que pueden ser aplicadas para alcanzar los resultados esperados.

Las siguientes cuatro etapas están relacionadas directamente con los datos. En la etapa de **Requerimientos de datos** se determinan las necesidades de contenido, formato y representaciones de los datos según el dominio de negocio y la aproximación analítica definida. En **Recolección de datos** se identifican y establecen las fuentes de datos relevantes al dominio del problema. A continuación, en la etapa de **Entendimiento de los datos**, se emplean elementos de estadística descriptiva y técnicas de visualización para entender el contenido de los datos, evaluar su calidad y descubrir los primeros *insights*. Por último en esta sección, en **Preparación de los datos** se realizan actividades de limpieza de datos, combinación con otras fuentes y transformación hacia nuevas variables o estructuras de datos útiles, con el fin de generar el conjunto de datos que será utilizado en la etapa de modelamiento.

En la etapa de **Modelamiento** se utiliza un subconjunto de los datos para desarrollar modelos descriptivos o predictivos usando la aproximación analítica previamente seleccionada.

En la etapa de **Evaluación** se revisan los resultados y la calidad del modelo para identificar si aborda adecuadamente el problema del negocio.

Luego de la aprobación del modelo seleccionado, se despliega en un entorno de producción o un entorno de pruebas equivalente (etapa de **Despliegue**). Finalmente, en la etapa de **Retroalimentación** se obtienen los resultados del modelo implementado para conocer su desempeño y evaluar cómo afecta el entorno de despliegue.

3.3. Marcos de referencia para analítica visual

Para el enfoque más específico del trabajo, el análisis visual, se utilizaron dos marcos de referencia. El primero está asociado al proceso de analítica visual presentado en (Keim Daniel, 2010), como se ilustra en la Figura 19. Este proceso contempla las etapas tradicionales de adquisición e integración de fuentes de datos heterogéneas, así como el preprocesamiento (limpieza, normalización, agrupamiento, integración) y transformación de los datos para ser explorados posteriormente. Una de las características de este proceso es el énfasis en dos maneras diferentes de abordar el análisis de datos: automatizado o visual.

El análisis automatizado contempla la aplicación de métodos de minería de datos para generar modelos de los datos. Una vez se ha generado el modelo se utiliza la visualización para interactuar con los datos, de tal manera que se logre evaluar y refinar el modelo a través de la modificación de parámetros o la selección de otros algoritmos de análisis. De otro lado, el análisis visual puede ser utilizado para la exploración de los datos y para confirmar las hipótesis generadas por los análisis automatizados (si éstos fueron generados previamente). La interacción del usuario con la visualización es necesaria para descubrir *insights* y, a su vez, estos conocimientos relevantes pueden ser utilizados para guiar la construcción de modelos en un análisis automatizado posterior.

Alternar entre los métodos automatizados y los visuales es característico del proceso de analítica visual y permite el refinamiento continuo y la verificación de resultados preliminares. Adicionalmente, el conocimiento puede ser obtenido desde la visualización, el análisis automatizado o desde las interacciones previas entre las visualizaciones, los modelos y los analistas del negocio (los usuarios).

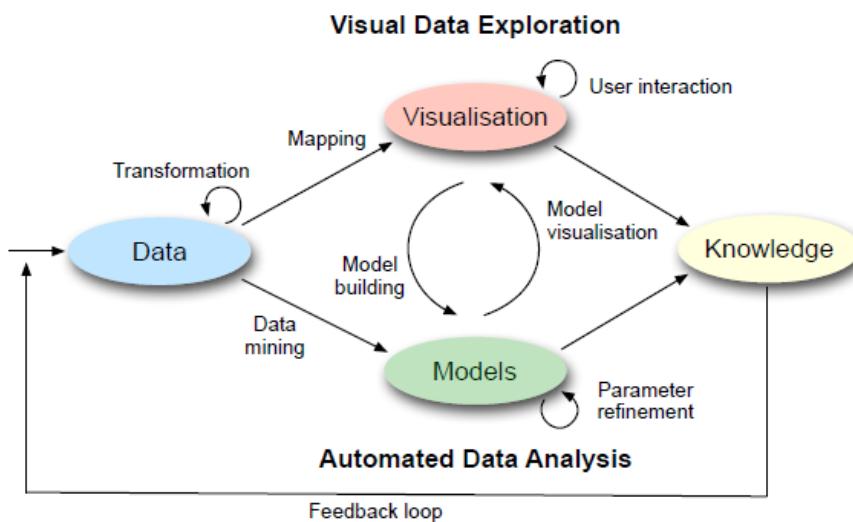


Figura 19. Proceso de analítica visual.

El segundo marco de referencia utilizado consiste en el framework de analítica visual de (Munzner, 2014), en el cual se plantean cuatro niveles fundamentales para el diseño de este tipo de soluciones: el dominio de negocio, la abstracción de los datos, la abstracción de las tareas de análisis y el establecimiento de los medios para visualizar los datos (Figura 20). Este framework asocia tres de estos niveles a tres preguntas fundamentales: qué puede visualizar el usuario (**what**, asociado a la abstracción de los datos, los conjuntos y

tipos de datos típicos), por qué se pretende utilizar una herramienta de visualización (**why**, asociado a la abstracción de las tareas de análisis, en términos de las acciones a realizar y el objetivo que se busca) y cómo la codificación visual y los modismos de interacción se utilizan para responder a dichas tareas (**how**, asociado a los medios para visualizar los datos, en términos de decisiones de diseño).

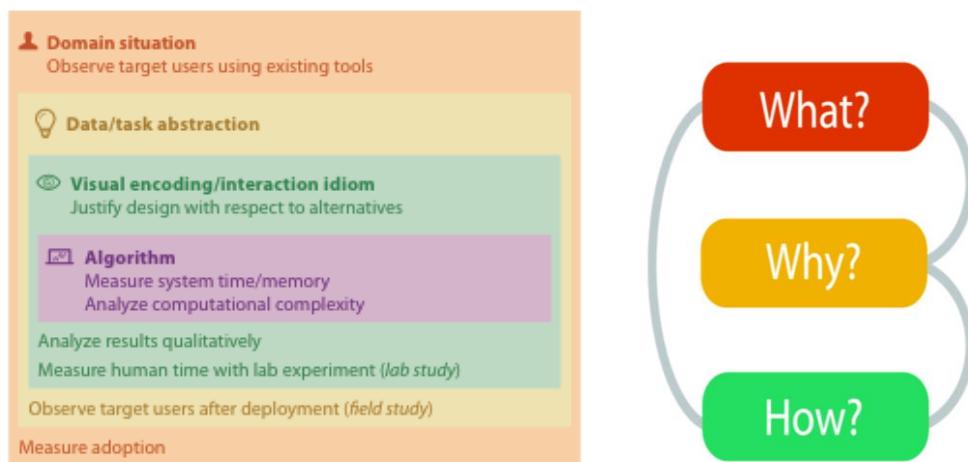


Figura 20. Framework para el diseño de visualizaciones.

3.4. Abordaje del proyecto

Tomando como base los marcos de referencia mencionados en las secciones anteriores, se establece un esquema para la planeación del trabajo de manera iterativa, considerando una fase por cada tipo de capacidad analítica (descriptiva y diagnóstica) y una serie de actividades basadas en los pasos de la metodología ASUM-DM (**¡Error! No se encuentra el origen de la referencia.**).

Para simplicidad del plan de trabajo, las etapas relacionadas con el tratamiento de los datos (requerimientos, recolección, entendimiento y preparación) se agruparon bajo una actividad llamada **Procesamiento de los Datos**.

Cada celda de la tabla (Fase/Actividad) representa un hito y tiene asociada una fecha según la planeación.

Tabla 2. Estructura iterativa del plan de trabajo a alto nivel basado en la metodología

ACTIVIDAD	FASE	
	ANÁLISIS DESCRIPTIVO	ANÁLISIS DIAGNÓSTICO
1. Entendimiento del negocio	t1	t1
2. Aproximación analítica	t2	t2
3. Procesamiento de Datos	t3	t3
4. Modelamiento y Visualización	t4	t4
5. Evaluación	t5	t5
6. Despliegue	t6	t6
7. Retroalimentación	t7	t7

En la ejecución de esta planeación se incorporan los pasos y elementos de los marcos de referencia para analítica visual, especialmente en la etapa de modelamiento, razón por la cual se nombra la actividad como **Modelamiento y Visualización**. En este punto vale la pena mencionar que las fuentes de información, las tareas de análisis según (Munzner, 2014) y las técnicas de procesamiento y visualización pueden ser diferentes en cada fase según los objetivos de uso de cada capacidad analítica. Esto se amplía con mayor detalle en la sección *Diseño de la visualización*, en el Capítulo 4. Resultados.

En relación con la aplicación y evaluación del trabajo propuesto, se utiliza como caso de estudio el conjunto de irregularidades de equipajes y vuelos con ubicación en la terminal 1 del Aeropuerto Internacional El Dorado, en la ciudad de Bogotá, en el periodo febrero 2018 – octubre 2019, para una aerolínea específica.

Capítulo 4. Resultados

Este capítulo presenta el desarrollo del trabajo y los resultados obtenidos, siguiendo el enfoque metodológico propuesto.

4.1. Fase 1: Análisis Descriptivo

4.1.1. Entendimiento del Negocio

En esta etapa se estudió tanto el proceso de manejo de equipajes como el de gestión de una irregularidad. El primero, fiel al proceso de industria descrito en capítulos anteriores de este documento y referenciado como el *bag journey*, se ilustra en la Figura 21.

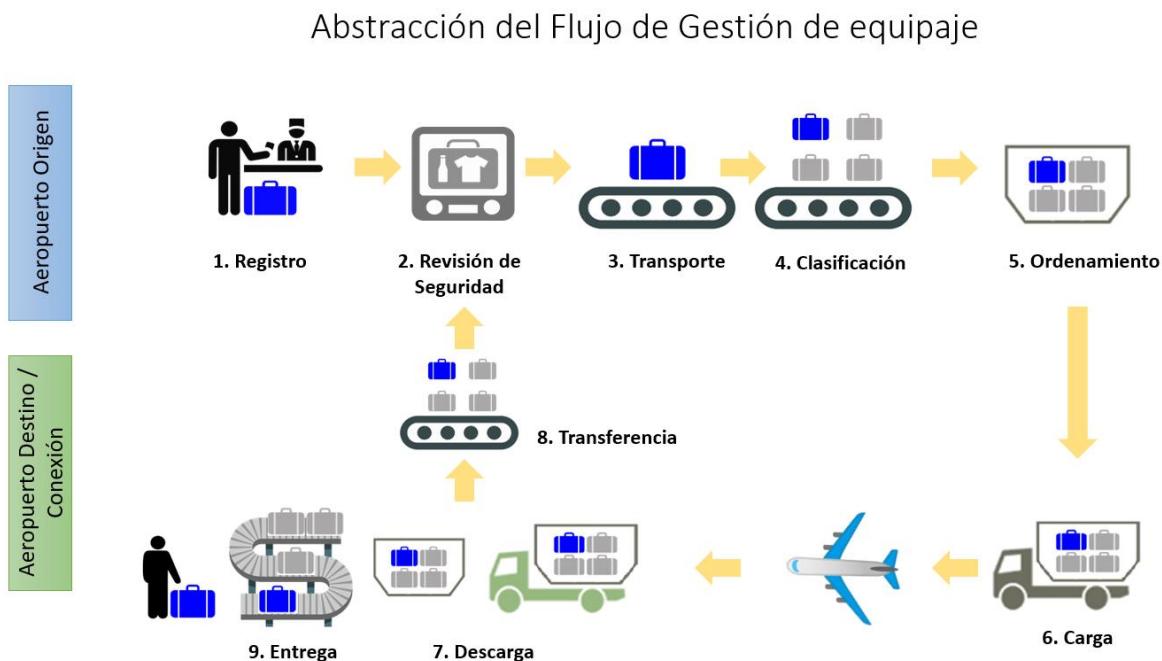


Figura 21. Abstracción del flujo de gestión de equipaje

Tabla 3. Descripción del flujo de gestión de equipajes

Paso	Ubicación	Actores	Actividades	Tiempos	Estado equipaje	Riesgos (tipos de incidente: Demora, Pérdida, Daño, Saqueo)
1. Registro	Counter aerolínea / punto de autoservicio	Asesor aerolínea	Impresión y asignación de etiqueta (<i>bag-tag</i>)	Tiempo de check-in	Registrado	Equipaje mal etiquetado: genera Demora
2. Revisión de seguridad	Punto establecido por el Aeropuerto	Personal seguridad aeropuerto	Revisión con scanner/caninos/manual	Tiempo de revisión (es)	Revisado	Tardanza en revisión / Hallazgo de contenido no permitido: genera Demora / Daño / Saqueo
3. Transporte	Bandas Aeropuerto	Automático	Traslado a zona de clasificación	Tiempo de tránsito banda	Revisado	Fallo en banda / caída: genera Demora, Pérdida.
4. Clasificación	Bandas asignadas a aerolínea	Personal aerolínea / sensor	Separación de equipaje por aerolínea y vuelo	Tiempo de clasificación	Asignado	Equipaje mal clasificado / asignación a otro vuelo-aerolínea: genera Demora
5. Ordenamiento	Bandas asignadas a aerolínea	Personal aerolínea	Organización de equipaje por vuelo, clase, prioridad, tamaño y peso	Tiempo de ordenamiento	Asignado	Asignación de vuelo equivocada / Manipulación equipaje: genera Demora, Daño
6. Carga	Bandas asignadas a aerolínea	Personal aerolínea	Traslado de equipaje a vehículos y luego a aeronave	Tiempo de carga	Cargado	Manipulación equipaje: genera Daño / Saqueo
7. Descarga	Bandas asignadas a aerolínea y vuelo	Personal aerolínea	Traslado de equipaje de aeronave a vehículos y luego a bandas de descarga	Tiempo de tránsito vehículo y descarga en banda	Descargado	Asignación de banda equivocada / Manipulación equipaje: genera Demora, Pérdida, Daño
8. Transferencia	Zona de entrega del aeropuerto de conexión	Personal aeropuerto conexión	Descargar equipaje. Ubicar en banda de conexión	Tiempo de traslado a sitio de conexión	Transferido	Demora en vuelo de origen o conexión / Equipaje no llega / llega dañado o incompleto: Demora, Pérdida, Daño
9. Entrega	Zona de entrega del aeropuerto	Pasajero	Reclamar equipaje. Radicar incidente si se presenta.	Tiempo de localización	Entregado	Demora en asignación de banda / Equipaje tarda en llegar / no llega / llega dañado o incompleto: Demora, Pérdida, Daño, Saqueo

Durante la revisión de cada uno de estos pasos se identifican los actores, las actividades involucradas, el sitio en que se llevan a cabo, los tiempos relevantes, el estado del equipaje y las situaciones que generan riesgos de que ocurra una irregularidad. Esta información se presenta en la Tabla 3.

De otro lado, cuando se presenta una irregularidad, se sigue un proceso que involucra a varios actores y puede variar de una aerolínea a otra. Esta participación múltiple se da por la naturaleza de esta industria en la que el personal de aerolíneas y aeropuertos son diferentes, también por la participación de diferentes empresas cuando se trata de vuelos en conexión o vuelos que son comercializados por una aerolínea y operados por otra. Adicionalmente, es usual encontrar en los aeropuertos una oficina encargada de la gestión de equipajes e irregularidades, que atiende de manera integral y centralizada los casos registrados en dicho lugar y que pueden provenir de cualquier otro aeropuerto o involucrar diversas aerolíneas.

La Figura 22 muestra en resumen el proceso de gestión de una irregularidad, el cual se describe con mayor detalle en la Tabla 4.

Esta representación general del flujo, como se mencionó anteriormente, puede variar en roles, pasos y actividades de una empresa a otra. Adicionalmente, pueden existir flujos de proceso más detallados por cada tipo de irregularidad (demora, pérdida, daño o saqueo), conformando todo un macroproceso.

En términos de desempeño, el proceso se mide a través del indicador MB (como se describió en la sección 1.1.4 *Indicadores del proceso*), no obstante se puede utilizar el costo de compensación, el tiempo de respuesta de una reclamación y la percepción del servicio como variables adicionales para su evaluación.

Resumen gestión de irregularidad

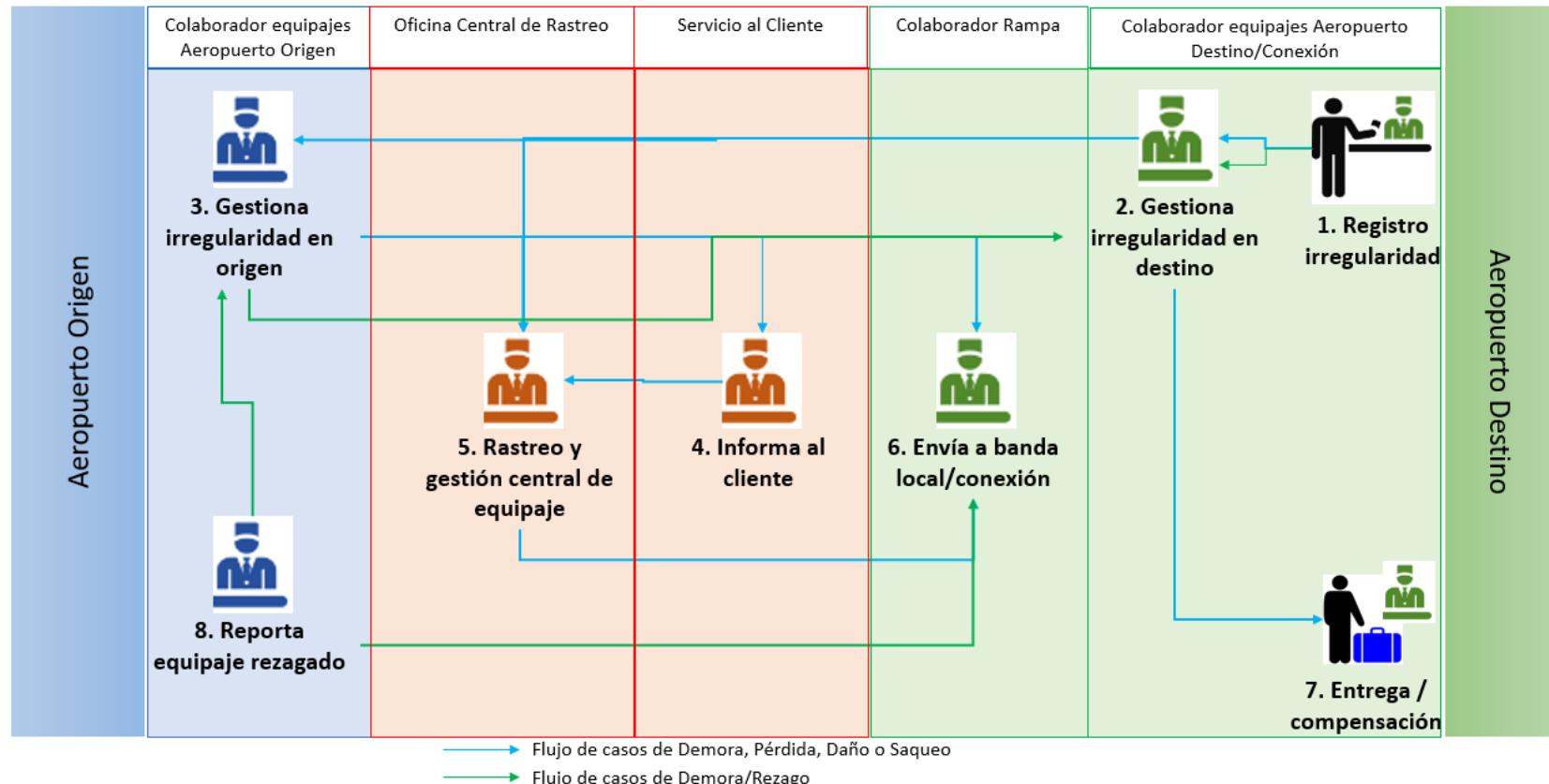


Figura 22. Flujo de gestión de una irregularidad

Tabla 4. Descripción del flujo de gestión de una irregularidad

Paso	Ubicación	Actores	Actividades	Tiempos
1. Registro irregularidad	Punto de atención aeropuerto destino	Colaborador equipajes	Crea caso. Toma evidencia (si es daño).	Tiempo de registro
2. Gestiona irregularidad en destino	Aeropuerto destino/conexión	Colaborador equipajes	Búsqueda equipaje, investigación del caso, verifica daño preexistente, ofrece compensación, verifica peso, revisa contenido (si es saqueo), escala con aeropuerto origen u oficina de rastreo.	Tiempo de gestión
3. Gestiona irregularidad en origen	Aeropuerto origen	Colaborador equipajes	Búsqueda de equipaje, etiquetar equipaje rezagado, enviar a bodega, ubicar vuelo para envío de rezago, revisar sobrantes, escalar con oficina de rastreo.	Tiempo de gestión
4. Informa al cliente	Call Center/Aeropuerto	Asesor servicio al cliente	Responder solicitudes de cliente, informar del estado del caso.	Tiempo de atención de llamada/consulta en sistemas
5. Rastreo y gestión central de equipaje	Aeropuerto BOG	Personal oficina	Recepción casos escalados desde estaciones luego de 5 días, búsqueda primaria y secundaria de equipaje, informe a CallCenter luego de 21 días para compensación, seguimiento a la entrega, donar equipaje luego de 6 meses, cierre de casos.	Tiempo de gestión
6. Envía a banda local/conexión	Aeropuerto destino/conexión	Colaborador Rampa	Coordinación con personal de seguridad, cargue y descargue de equipaje, manejo de equipaje con prioridad, envío a vuelo de conexión, revisión entrega en banda adecuada.	Tiempos de carga Tiempos de descarga Tiempo de traslado de equipaje
7. Entrega / compensación	Aeropuerto destino/conexión	Colaborador equipajes	Entregar equipaje rezagado, ofrecer compensación o reparación.	Tiempo de gestión
8. Reporta equipaje rezagado	Aeropuerto origen	Colaborador equipajes	Identificar, registrar y reportar equipaje rezagado.	Tiempo de registro y gestión

4.1.2. Aproximación analítica

De acuerdo con el proceso de analítica visual de (Keim Daniel, 2010), referenciado en la sección 3.3, se utilizaron tanto los métodos automatizados como la aproximación visual. Para la primera parte se utilizaron estadísticos básicos para el entendimiento de las irregularidades y se realizó un perfilamiento de los datos, dentro del cual se aplican de manera automática algunas correlaciones preliminares (Pearson, Spearman) y se generan histogramas que apoyan el análisis descriptivo.

Para la segunda parte se diseñó un tablero de control conformado por diferentes visualizaciones que permiten realizar una exploración de los datos de irregularidades con el fin de lograr un entendimiento de cuántos, cuándo, dónde se presentaron y qué causas se les atribuyeron.

4.1.3. Procesamiento de Datos

En esta etapa se identifica como principal requerimiento de datos aquel conjunto que contiene las irregularidades presentadas, dentro del alcance del caso de estudio indicado en la sección 2.3.

En la etapa de recolección de los datos se identifica la necesidad de incluir otros conjuntos adicionales para complementar los atributos de las irregularidades, como el conjunto de países, aeropuertos y códigos de causas (RL). Adicionalmente se identifica la necesidad fundamental de contar con el número de pasajeros transportados para tener la capacidad de calcular el indicador MB (como se describe en la sección 1.1.4) y poder analizar su comportamiento. Para esto se incorpora un dataset de vuelos que contiene los atributos de número de pasajeros y número de maletas transportadas.

El modelo de datos en esta etapa está conformado por los conjuntos de datos de irregularidades, vuelos, aeropuertos, países y causas de irregularidades (RL), como se muestra en la Figura 23.

Las relaciones están dadas por la siguiente lógica:

- A una irregularidad se asigna un solo aeropuerto responsable.
- En un aeropuerto se pueden presentar muchas irregularidades.

- Una irregularidad puede tener una sola causa (RL) y una causa puede ser asignada a muchas irregularidades.
- A una irregularidad se asigna un solo vuelo responsable y un vuelo puede ser causante de muchas irregularidades.
- Un aeropuerto opera muchos vuelos y un vuelo está relacionado con al menos un aeropuerto de origen y uno de destino.
- Un aeropuerto pertenece a un país y un país puede tener varios aeropuertos.

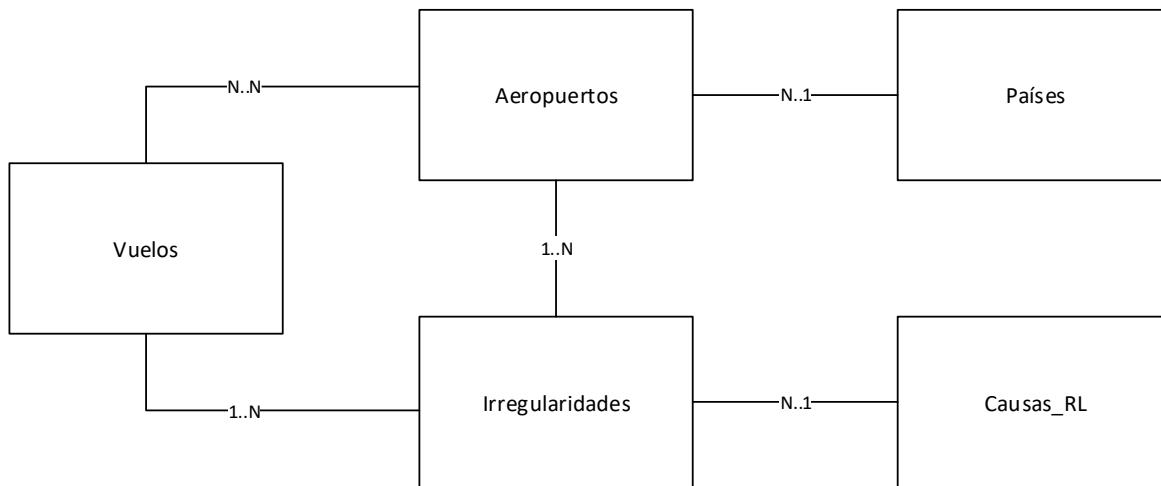


Figura 23. Modelo de datos en el análisis descriptivo

La Tabla 5 y la Tabla 6 resumen a nivel conceptual el entendimiento de los datos más relevantes en esta etapa.

Tabla 5. Conjunto de datos de irregularidades

Atributo	Tipo	Descripción
Fecha de creación	Fecha-hora	Fecha en que se radica la irregularidad.
Lugar de creación	Alfanumérico	Aeropuerto en que se radica la reclamación. Hace parte de una jerarquía a la cual pertenece tanto el aeropuerto como la región de este (usualmente en términos del continente, conjunto de países o zona donde opera la aerolínea).
Código de la irregularidad	Alfanumérico	Código único asignado por el sistema.
Lugar de asignación de responsabilidad	Alfanumérico	Aeropuerto al que se asigna la responsabilidad de la irregularidad. Hace parte de la misma jerarquía de ubicación (región, continente, conjunto de países o zona donde opera la aerolínea).

Atributo	Tipo	Descripción
Tipo de Causa	Alfanumérico	Conjunto de códigos primarios según la tabla de códigos RL descrita en la sección 1.1.1., Figura 3.
Código de Causa (RL)	Alfanumérico	Conjunto de códigos secundarios según la tabla de códigos RL descrita en la sección 1.1.1., Figura 3.
Tipo de Irregularidad	Alfanumérico	Clasificación de la irregularidad entre: demora, pérdida, daño o saqueo.
Estado	Alfanumérico	Indica si el caso se encuentra abierto o cerrado.
Ruta	Alfanumérico	Secuencia de segmentos y destinos por los que transitó el pasajero que radicó la irregularidad.
Vuelo afectado	Alfanumérico	Código del vuelo al cual se asoció la responsabilidad de la irregularidad.
Fecha del vuelo afectado	Fecha	Fecha programada del vuelo al cual se asoció la responsabilidad de la irregularidad.
Empresa de la ruta	Alfanumérico	Empresa o línea aérea que comercializa la ruta.
Empresa transportadora	Alfanumérico	Empresa o línea aérea que opera la ruta.
Maletas afectadas	Numérico	Número de maletas afectadas según se indica en la reclamación.
Costo	Numérico	Valor pagado por compensación al pasajero.

Tabla 6. Conjunto de datos de vuelos

Atributo	Tipo	Descripción
Número de vuelo	Numérico	Identificador del vuelo
Fecha de salida	Fecha-hora	Fecha programada del vuelo
Origen	Alfabético	Código del aeropuerto de origen
Destino	Alfabético	Código del aeropuerto de destino
Pasajeros	Numérico	Número de pasajeros transportados
Maletas	Numérico	Número de maletas transportadas
Tipo de Ruta	Alfabético	Indica si el vuelo es Doméstico o Internacional

En el caso de los aeropuertos, países y códigos de causa, los atributos están conformados como pares código-nombre, por simplicidad no se amplían en esta sección.

En términos de calidad de los datos se identifican faltantes especialmente en los atributos de fecha y vuelo responsable (en el 12.4% de los datos) y en el costo de compensación del caso (7.5%). También se identifica un conjunto bajo de casos abiertos y otro más sin asignación de ubicación responsable. Se comprende por la dinámica de la gestión de

irregularidades que el proceso de investigación de la causa toma tiempos que pueden depender de la complejidad del caso, también que durante este tiempo estos datos no tendrán asignación de responsabilidad.

En términos de preparación de los datos, se llevan a cabo tareas adicionales de estandarización de códigos, estandarización y asignación de valores nulos, identificación de duplicados e identificación de atributos sin valor para el análisis (o con pésima calidad en términos de completitud). Tareas de conversión de formato y tipos de dato son necesarias para los atributos de tipo fecha y fecha-hora. Así mismo se analiza la consistencia entre algunos atributos que guardan relación, como los códigos de causa (RL) y el tipo de irregularidad asignado (demora, pérdida, daño o saqueo).

Con respecto al volumen de datos, las cifras obtenidas a partir del análisis se presentan en la Tabla 7. El conjunto de datos de mayor peso es el de irregularidades, tanto por el número de ítems como de atributos.

Tabla 7. Volumen de datos

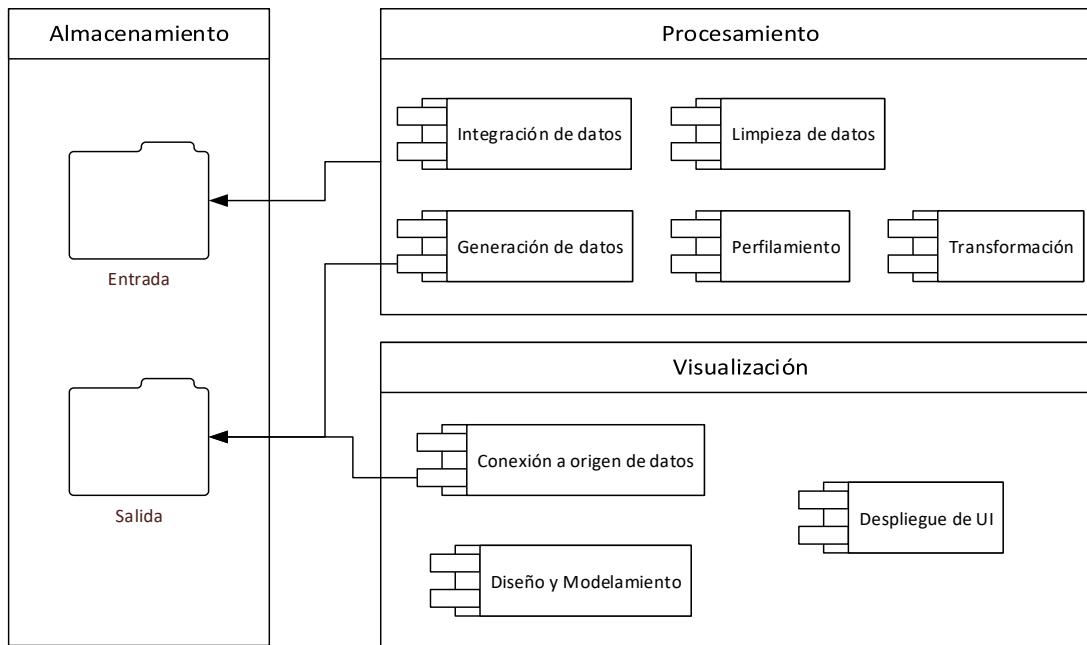
Conjunto de datos	Ítems	Atributos	Tamaño
Irregularidades	390.203	28	84 MB
Vuelos	447.000	13	38 MB
Aeropuertos	9.000	13	1 MB
Países	252	5	88 KB
Causas_RL	48	5	2 KB

4.1.4. Modelamiento y Visualización

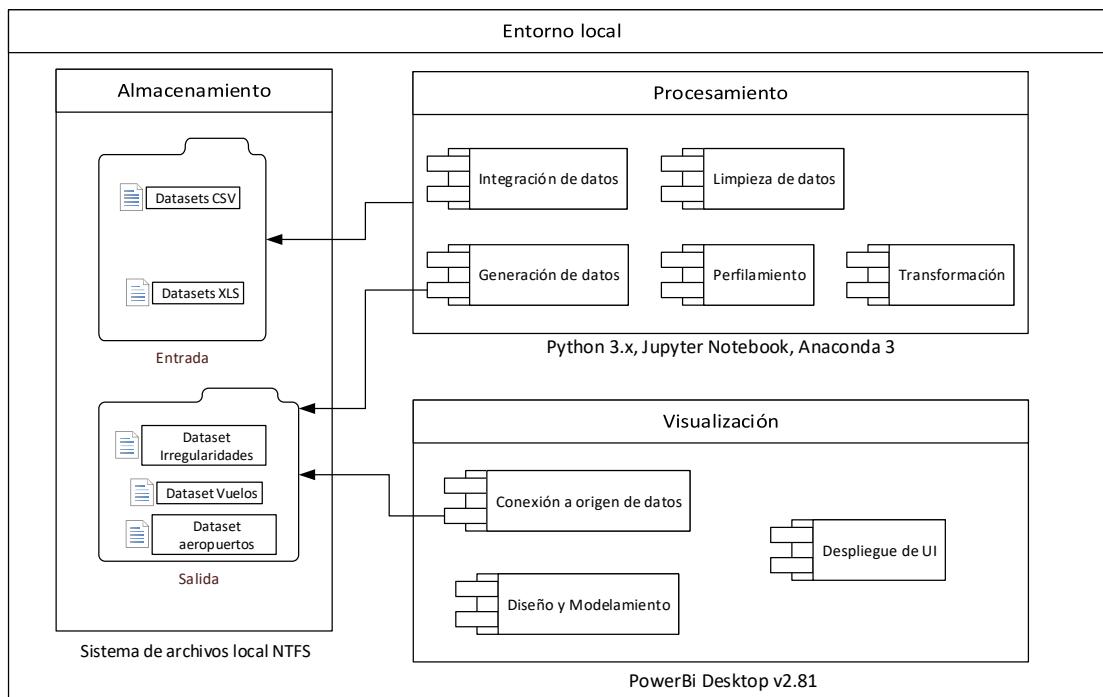
Esta sección presenta los temas relacionados con el diseño de la herramienta, los análisis automatizados sobre los datos y el diseño de la herramienta de visualización.

4.1.4.1. Diseño de la herramienta

El diseño de alto nivel se representa en la vista lógica de la Figura 24. Una capa lógica de almacenamiento es la encargada de la persistencia de los datos tanto de entrada como de salida. La capa de procesamiento es la encargada de la integración, limpieza, perfilamiento y transformación de datos. También genera nuevos conjuntos de datos previo a la etapa de modelamiento. En la capa de visualización se realiza el diseño e implementación de la herramienta de análisis visual que será desplegada al usuario final.

**Figura 24. Vista lógica de la solución**

En esta fase del proyecto la implementación completa se realiza en un entorno local, como se muestra en la vista física de la Figura 25.



Laptop Windows 10 - Intel Core i5 - 2vCPU 2.5 GHz - 8GB RAM - 200 GB HDD

Figura 25. Vista física de la solución

El almacenamiento se realiza en el sistema de archivos del equipo portátil utilizado. Para el procesamiento se utiliza el entorno Anaconda3, que contiene tanto el lenguaje Python 3.x como el entorno de desarrollo de Jupyter Notebook. La visualización se realiza con la herramienta PowerBi versión 2.81 de escritorio, y utiliza como origen de datos los archivos generados en la capa de procesamiento.

4.1.4.2. Análisis asistido¹

En esta actividad se aplican técnicas de perfilamiento que generan estadísticos básicos sobre los datos. Utilizando librerías y funciones de Python se realiza la exploración del conjunto de datos de irregularidades, se descubre su distribución y tendencia a través del año, se identifican los aeropuertos donde se presenta el mayor número de casos, así como las causas más frecuentes, entre otros. Parte de los hallazgos identificados en esta etapa se presentan en la Tabla 8.

Tabla 8. Ejemplo de hallazgos del análisis descriptivo asistido para 2018

Descripción
<ul style="list-style-type: none"> • El 5.7% de los casos se encuentran aún abiertos, a pesar de ser de 2018. • La irregularidad más común es Demoras, con el 83.7% de los casos. Le sigue Daños con 14.1%, Saqueos con 2% y Pérdidas con 0.3%. • El top 5 de tipos de causa (RL) es: <ul style="list-style-type: none"> • Aeropuerto en general 21.9% • Aeropuerto de conexión 18.6% • Aeropuerto origen - equipaje no abordado 17.5% • Danos 14.1% • Misceláneos 11.7% • El top 3 de códigos RL asignados es: <ul style="list-style-type: none"> • 64 (Problemas con instalación aeroportuaria/Problemas con el sistema en bandas de selección) 14.5%, • 26 (Equipaje rezagado por restricciones de peso o espacio) 12.4% y • 86 (Daño en equipaje vuelo origen afecta funcionalidad) 12.1%. • El 47.5% de los casos fueron creados al día siguiente del vuelo, y el 38.6% de los casos fueron creados el mismo día. • Las 3 estaciones que más casos crean son BOG, CLO y MED (8.7, 8.7 y 7.4%), sin embargo se asigna la responsabilidad de la irregularidad a BOG en el 56.5% de los casos, seguido por MIA (4.7%), SAL (3.7%) y MAD (3.5%). • El mayor costo de compensación es de BOG, MIA, MAD, LIM y JFK. • Los meses de mayor reclamo son dic y ene, seguidos por jun, ago, nov y jul. • No hay ruta frecuente, requiere mayor análisis. • No hay correlaciones relevantes entre los datos numéricos iniciales.

¹ Aunque en la terminología de (Keim Daniel, 2010) se usa el término automatizado, se prefiere el uso del término ‘asistido’, aunque muchas de estas tareas se apoyan típicamente en scripts y funciones de lenguajes especializados como Python o R, es decisión del analista qué tipo de función utilizar, en qué momento y cuáles de las salidas son relevantes, según el tipo de proyecto y la etapa en la que se encuentre.

Adicionalmente, se incorporan los datos de vuelos, pasajeros y maletas, y se realizan algunas tareas iniciales para analizar la correlación entre algunas de estas variables (Figura 26). En la imagen se puede ver la relación proporcional entre pasajeros y maletas, entre las irregularidades de demora presentadas y el número de pasajeros y maletas, y en menor grado también con los casos de daños presentados.

	PAX_VUELO	BAGS_VUELO	DEMORA	PERDIDA	DANO	SAQUEO	TOTAL_CASOS	RUTA
PAX_VUELO	1.000000	0.957073	0.983851	0.552898	0.888454	0.623855	0.986322	0.011048
BAGS_VUELO	0.957073	1.000000	0.964976	0.639584	0.878904	0.701409	0.968932	0.095636
DEMORA	0.983851	0.964976	1.000000	0.574341	0.854731	0.622496	0.999253	0.024176
PERDIDA	0.552898	0.639584	0.574341	1.000000	0.579160	0.631752	0.583387	0.191301
DANO	0.888454	0.878904	0.854731	0.579160	1.000000	0.724018	0.873776	0.055948
SAQUEO	0.623855	0.701409	0.622496	0.631752	0.724018	1.000000	0.641177	0.179505
TOTAL_CASOS	0.986322	0.968932	0.999253	0.583387	0.873776	0.641177	1.000000	0.028378
RUTA	0.011048	0.095636	0.024176	0.191301	0.055948	0.179505	0.028378	1.000000

Figura 26. Análisis asistido - correlación entre atributos numéricos

Al realizar este mismo análisis filtrando los datos para una región específica, se identifican algunos comportamientos ligeramente diferentes. Por ejemplo, al pasar a la región europea los valores de correlación entre los casos de pérdida, daño y saqueo contra el número de pasajeros y maletas aumentan. Sin embargo, en el caso de las irregularidades de demora el valor se reduce. Este y otros tipos de consultas generan hipótesis que luego se analizan de una manera más ágil y práctica con la solución de analítica visual.

4.1.4.3. Diseño de la visualización

En las secciones a continuación se utiliza como referente el framework de (Munzner, 2014) mencionado en el numeral 3.3 (Marcos de referencia para analítica visual). El nivel correspondiente al dominio de negocio será omitido por cuanto ya se abarcó en la sección 4.1.1 (Entendimiento del negocio), con excepción de la identificación del usuario de la visualización, el cual será mencionado en el apartado de *abstracción de las tareas*.

ABSTRACCIÓN DE LOS DATOS

El tipo de conjunto de datos presente en toda la información recolectada es el de tablas, se cuenta con una para cada entidad representada en el modelo de la Figura 23. La disponibilidad de cada conjunto es estática, ya que proviene de archivos planos y no de una fuente en línea que se actualice constantemente.

Los tipos de atributos identificados en este punto se describen a continuación:

Tabla 9. Tipos de atributos de las irregularidades

Atributo	Tipo	Dirección / Ordenamiento	Jerarquía
Fecha de creación	Ordinal (usado como fecha) Cuantitativo (usado como día, mes y año por separado)	Cíclico	Si (Año / Mes / Día)
Lugar de creación	Categórico	NA	Si (Región, País, Ciudad, Aeropuerto)
Código de la irregularidad	Categórico	NA	No
Lugar de asignación de responsabilidad	Categórico	NA	Si (Región, País, Ciudad, Aeropuerto)
Tipo de Causa	Cuantitativo	Secuencial	Si (códigos primarios de RL)
Código de Causa (RL)	Cuantitativo	Secuencial	Si (códigos secundarios de RL)
Tipo de Irregularidad	Categórico	NA	NA
Estado	Categórico	NA	NA
Ruta	Categórico	NA	NA
Vuelo afectado	Cuantitativo	Secuencial	NA
Fecha del vuelo afectado	Ordinal (usado como fecha) Cuantitativo (usado como día, mes y año por separado)	Cíclico	Si (Año / Mes / Día)
Empresa comercial de la ruta	Categórico	NA	NA
Empresa transportadora	Categórico	NA	NA
Maletas afectadas	Cuantitativo	Secuencial	NA
Costo	Cuantitativo	Secuencial	NA

Tabla 10. Tipos de atributos de los vuelos

Atributo	Tipo	Dirección / Ordenamiento	Jerarquía
Número de vuelo	Cuantitativo	Secuencial	NA
Fecha de salida	Ordinal (usado como fecha) Cuantitativo (usado como día, mes y año por separado)	Cíclico	Si (Año / Mes / Día)
Origen	Categórico	NA	No
Destino	Categórico	NA	No
Pasajeros	Cuantitativo	Secuencial	NA
Maletas	Cuantitativo	Secuencial	NA
Tipo de Ruta	Categórico	NA	No

ABSTRACCIÓN DE LAS TAREAS

Antes de la descripción de las acciones y los objetivos, es necesario identificar el perfil del usuario de negocio que realizará estas tareas y se beneficiará con la solución. En línea con el proceso de negocio descrito en la sección 4.1.1, específicamente el flujo de gestión de irregularidades, existe un **Coordinador de la Oficina Central de Rastreo de Equipajes**, localizado en el aeropuerto. Adicionalmente se considera un **Gerente del Departamento de Gestión de Equipajes**, por parte de las aerolíneas, como otro actor al cual estará destinada esta solución. No se realizará distinción entre las tareas asignadas a uno u otro actor, se entiende que ambos tienen los mismos objetivos de negocio: analizar el comportamiento de las irregularidades, identificar sus causas y elaborar planes de acción para reducirlas.

Respecto de las tareas identificadas para estos actores, descritas en términos de **acciones** y **objetivos**, a continuación se enumeran las que corresponden a esta fase del proyecto (análisis descriptivo):

- T1. **Presentar la distribución** de los tipos de irregularidades y del indicador MB a través del año.
- T2. **Identificar la tendencia** de las irregularidades, los costos y el indicador MB a lo largo del año.
- T3. **Presentar la correlación** entre los tipos de irregularidad y el número de irregularidades, costo e indicador MB.
- T4. **Presentar la correlación** entre las irregularidades presentadas por región y el indicador MB, el número de casos y el costo.
- T5. **Buscar los extremos** que representan los aeropuertos responsables con el mayor costo, indicador MB y número de casos.
- T6. **Buscar los extremos** que representan los códigos de causas (RL) asignadas con la mayor frecuencia según el tipo de irregularidad y la región.
- T7. **Identificar valores atípicos** en el número de irregularidades, el indicador MB o los costos en momentos específicos del año.

DECISIONES DE DISEÑO

A continuación se describen las consideraciones de diseño por cada tarea identificada, incluyendo los modismos de visualización, canales, marcas, decisiones y observaciones

(Tabla 11). La Figura 27 presenta de manera general la vista completa de la solución en esta fase.

Tabla 11. Decisiones de diseño para la visualización

Representación	Codificación	Tarea
<p>Cifras por Región</p>	<p>Modismo: <i>BubbleChart</i> Marcas: Puntos. Canales:</p> <ul style="list-style-type: none"> Color, para identificar la región. Posición en X, para representar el costo. Posición en Y, para representar el indicador MB. Tamaño, para reflejar el volumen de los casos. <p>Decisiones de diseño:</p> <ul style="list-style-type: none"> Se usan los atributos de IndMB y Costo, ya que son los principales motivadores de los planes de acción en el manejo de irregularidades. 	T4
<p>Distribución de casos y MB por periodo</p> <p>Tipo ● Daños ● Demoras ● Pérdidas ● Saqueos ● IndMB</p>	<p>Modismo: <i>Stacked Barchart</i> Marcas: Líneas. Canales:</p> <ul style="list-style-type: none"> Longitud, para indicar el número de casos presentados. Color, para identificar el tipo de irregularidad. Posición en X, para representar cada ítem de la variable tiempo (mes, día). Orden, de acuerdo con la variable temporal (meses, días) <p>Decisiones de diseño:</p> <ul style="list-style-type: none"> Este gráfico incorpora un atributo adicional para el Indicador MB representado con el modismo de <i>Linechart</i>, usando la escala de medición vertical de la derecha, con el fin de poder comparar la tendencia en el comportamiento del indicador con la tendencia en el número de casos presentados. Se utilizan barras apiladas y no barras paralelas para que el diagrama pueda escalar mejor cuando se amplía el detalle en la variable de tiempo. Se asume que la tarea de identificar la tendencia de las irregularidades en el tiempo es más importante que la comparación de la tendencia entre cada tipo de irregularidad en particular. No obstante, las barras apiladas permiten identificar de manera general la predominancia de al menos dos tipos de irregularidades sobre las demás a lo largo del año. 	T1, T2, T7

Representación	Codificación	Tarea																																																																		
<p>Cifras por Tipo de Irregularidad</p> <table border="1"> <thead> <tr> <th>Tipo de Irregularidad</th> <th>IndMB</th> <th>Costo</th> </tr> </thead> <tbody> <tr> <td>Demoras</td> <td>~18</td> <td>~\$ 1.5M</td> </tr> <tr> <td>Daños</td> <td>~8</td> <td>~\$ 1.2M</td> </tr> <tr> <td>Perdidas</td> <td>~4</td> <td>~\$ 0.5M</td> </tr> </tbody> </table>	Tipo de Irregularidad	IndMB	Costo	Demoras	~18	~\$ 1.5M	Daños	~8	~\$ 1.2M	Perdidas	~4	~\$ 0.5M	<p>Modismo: BubbleChart Marcas: Puntos. Canales:</p> <ul style="list-style-type: none"> Color, para identificar el tipo de irregularidad. Posición en X, para representar el costo. Posición en Y, para representar el indicador MB. Tamaño, para reflejar el número de casos. <p>Decisiones de diseño:</p> <ul style="list-style-type: none"> Se usan los atributos de IndMB y Costo, ya que son los principales motivadores de los planes de acción en el manejo de irregularidades. Se mantiene el color usado para cada tipo de irregularidad en todo el diseño, por uniformidad y para facilitar la identificación por parte del usuario. Se usan etiquetas adicionales al ubicarse sobre cada punto (tipo de irregularidad), con los valores y porcentajes de costo y número de casos, para poder contar con los valores exactos en cada caso. 	T3																																																						
Tipo de Irregularidad	IndMB	Costo																																																																		
Demoras	~18	~\$ 1.5M																																																																		
Daños	~8	~\$ 1.2M																																																																		
Perdidas	~4	~\$ 0.5M																																																																		
<p>Top aeropuertos (MB)</p> <table border="1"> <thead> <tr> <th>Aeropuerto</th> <th>IndMB</th> </tr> </thead> <tbody> <tr><td>IBE</td><td>74,42</td></tr> <tr><td>CUZ</td><td>71,88</td></tr> <tr><td>PPN</td><td>70,61</td></tr> <tr><td>FLA</td><td>58,82</td></tr> <tr><td>LPB</td><td>44,58</td></tr> <tr><td>MZL</td><td>33,91</td></tr> <tr><td>VVC</td><td>29,86</td></tr> <tr><td>JUL</td><td>29,81</td></tr> <tr><td>TCO</td><td>20,96</td></tr> <tr><td>BOG</td><td>18,33</td></tr> </tbody> </table> <p>Top aeropuertos (Costo)</p> <table border="1"> <thead> <tr> <th>Aeropuerto</th> <th>Costo</th> </tr> </thead> <tbody> <tr><td>BOG</td><td>\$ 2.316.567</td></tr> <tr><td>LIM</td><td>\$ 127.860</td></tr> <tr><td>CUZ</td><td>\$ 89.021</td></tr> <tr><td>CLO</td><td>\$ 75.056</td></tr> <tr><td>SCL</td><td>\$ 63.747</td></tr> <tr><td>GRU</td><td>\$ 61.771</td></tr> <tr><td>MDE</td><td>\$ 57.867</td></tr> <tr><td>CTG</td><td>\$ 37.790</td></tr> <tr><td>BAQ</td><td>\$ 28.285</td></tr> <tr><td>LPB</td><td>\$ 26.051</td></tr> </tbody> </table> <p>Top aeropuertos (casos)</p> <table border="1"> <thead> <tr> <th>Aeropuerto</th> <th>Casos</th> </tr> </thead> <tbody> <tr><td>BOG</td><td>100906</td></tr> <tr><td>CUZ</td><td>4037</td></tr> <tr><td>LIM</td><td>3941</td></tr> <tr><td>CLO</td><td>2142</td></tr> <tr><td>SCL</td><td>1992</td></tr> <tr><td>MDE</td><td>1927</td></tr> <tr><td>GRU</td><td>1358</td></tr> <tr><td>CTG</td><td>1022</td></tr> <tr><td>LPB</td><td>796</td></tr> <tr><td>BAQ</td><td>726</td></tr> </tbody> </table>	Aeropuerto	IndMB	IBE	74,42	CUZ	71,88	PPN	70,61	FLA	58,82	LPB	44,58	MZL	33,91	VVC	29,86	JUL	29,81	TCO	20,96	BOG	18,33	Aeropuerto	Costo	BOG	\$ 2.316.567	LIM	\$ 127.860	CUZ	\$ 89.021	CLO	\$ 75.056	SCL	\$ 63.747	GRU	\$ 61.771	MDE	\$ 57.867	CTG	\$ 37.790	BAQ	\$ 28.285	LPB	\$ 26.051	Aeropuerto	Casos	BOG	100906	CUZ	4037	LIM	3941	CLO	2142	SCL	1992	MDE	1927	GRU	1358	CTG	1022	LPB	796	BAQ	726	<p>Modismo: Horizontal Bar Charts Marcas: Líneas. Canales:</p> <ul style="list-style-type: none"> Longitud, para representar el valor del atributo en cada diagrama (indicador MB, costo y número de casos, de izq. a der.). Orden, cada diagrama está ordenado por el valor del atributo correspondiente de manera descendente. <p>Decisiones de diseño:</p> <ul style="list-style-type: none"> Se usa un diagrama para cada atributo: IndMB, Costo y Número de Casos, ya que son los principales motivadores de los planes de acción en el manejo de irregularidades. Aunque no es necesario para la tarea indicada, por uniformidad se usa el mismo color para cada uno de estos atributos en todo el diseño. 	T5
Aeropuerto	IndMB																																																																			
IBE	74,42																																																																			
CUZ	71,88																																																																			
PPN	70,61																																																																			
FLA	58,82																																																																			
LPB	44,58																																																																			
MZL	33,91																																																																			
VVC	29,86																																																																			
JUL	29,81																																																																			
TCO	20,96																																																																			
BOG	18,33																																																																			
Aeropuerto	Costo																																																																			
BOG	\$ 2.316.567																																																																			
LIM	\$ 127.860																																																																			
CUZ	\$ 89.021																																																																			
CLO	\$ 75.056																																																																			
SCL	\$ 63.747																																																																			
GRU	\$ 61.771																																																																			
MDE	\$ 57.867																																																																			
CTG	\$ 37.790																																																																			
BAQ	\$ 28.285																																																																			
LPB	\$ 26.051																																																																			
Aeropuerto	Casos																																																																			
BOG	100906																																																																			
CUZ	4037																																																																			
LIM	3941																																																																			
CLO	2142																																																																			
SCL	1992																																																																			
MDE	1927																																																																			
GRU	1358																																																																			
CTG	1022																																																																			
LPB	796																																																																			
BAQ	726																																																																			

 <p>Casos por Tipo de Causa y Cod.Causa (RL)</p> <table border="1"> <thead> <tr> <th>Categoría</th> <th>Porcentaje</th> </tr> </thead> <tbody> <tr> <td>aeropuerto en general</td> <td>17.24%</td> </tr> <tr> <td>aeropuerto origen/destino</td> <td>14.06%</td> </tr> <tr> <td>danos 86-daño en equipaje</td> <td>10.97%</td> </tr> <tr> <td>error en el equipaje</td> <td>6.88%</td> </tr> <tr> <td>aeropuerto destino</td> <td>6.69%</td> </tr> <tr> <td>misceláneo</td> <td>5.88%</td> </tr> <tr> <td>errores en la documentación</td> <td>4.79%</td> </tr> <tr> <td>aeropuerto en vuelo</td> <td>4.76%</td> </tr> <tr> <td>otro tipo de error</td> <td>3.95%</td> </tr> <tr> <td>daños en el equipaje</td> <td>3.94%</td> </tr> <tr> <td>aeropuerto en vuelo</td> <td>3.61%</td> </tr> <tr> <td>misceláneo</td> <td>3.26%</td> </tr> <tr> <td>otro tipo de error</td> <td>3.47%</td> </tr> </tbody> </table>	Categoría	Porcentaje	aeropuerto en general	17.24%	aeropuerto origen/destino	14.06%	danos 86-daño en equipaje	10.97%	error en el equipaje	6.88%	aeropuerto destino	6.69%	misceláneo	5.88%	errores en la documentación	4.79%	aeropuerto en vuelo	4.76%	otro tipo de error	3.95%	daños en el equipaje	3.94%	aeropuerto en vuelo	3.61%	misceláneo	3.26%	otro tipo de error	3.47%	<p>Modismo: Tree Chart Marcas: Áreas. Canales:</p> <ul style="list-style-type: none"> Color, para identificar cada categoría de códigos de causa. Contención, para expresar la jerarquía entre los códigos primarios y los secundarios (RL) (los primarios contienen a los secundarios) Orden, descendente por número de casos presentados por cada causa. <p>Decisiones de diseño:</p> <ul style="list-style-type: none"> Se utiliza la proporción y no el número casos, ya que para la tarea de identificar la causa más frecuente un porcentaje es más fácil de comparar (escala de 1 a 100) que múltiples valores en una escala más amplia. Aunque para esta tarea un diagrama de barras horizontales funcionaría, la descripción de cada categoría representa una limitación de espacio. El TreeMap permite manejar mejor esta limitación, también escala mejor para representar más categorías. Se utilizan etiquetas adicionales al ubicarse sobre cada categoría para ver la descripción completa de cada causa, para compensar las restricciones de espacio. También los valores y porcentajes de costo y número de casos, así como el indicador MB, para poder identificar en qué grado la causa contribuye a cada uno de estos atributos. 	T6								
Categoría	Porcentaje																																					
aeropuerto en general	17.24%																																					
aeropuerto origen/destino	14.06%																																					
danos 86-daño en equipaje	10.97%																																					
error en el equipaje	6.88%																																					
aeropuerto destino	6.69%																																					
misceláneo	5.88%																																					
errores en la documentación	4.79%																																					
aeropuerto en vuelo	4.76%																																					
otro tipo de error	3.95%																																					
daños en el equipaje	3.94%																																					
aeropuerto en vuelo	3.61%																																					
misceláneo	3.26%																																					
otro tipo de error	3.47%																																					
 <p>Distribución casos y costos por periodo</p> <p>● Casos ● Costo</p> <table border="1"> <thead> <tr> <th>Mes</th> <th>Casos</th> <th>Costo</th> </tr> </thead> <tbody> <tr> <td>febrero</td> <td>~15K</td> <td>~\$300K</td> </tr> <tr> <td>marzo</td> <td>~19K</td> <td>~\$450K</td> </tr> <tr> <td>abril</td> <td>~17K</td> <td>~\$400K</td> </tr> <tr> <td>mayo</td> <td>~14K</td> <td>~\$350K</td> </tr> <tr> <td>junio</td> <td>~19K</td> <td>~\$450K</td> </tr> <tr> <td>julio</td> <td>~15K</td> <td>~\$400K</td> </tr> <tr> <td>agosto</td> <td>~16K</td> <td>~\$500K</td> </tr> <tr> <td>septiembre</td> <td>~13K</td> <td>~\$300K</td> </tr> <tr> <td>octubre</td> <td>~14K</td> <td>~\$300K</td> </tr> <tr> <td>noviembre</td> <td>~15K</td> <td>~\$350K</td> </tr> <tr> <td>diciembre</td> <td>~22K</td> <td>~\$500K</td> </tr> </tbody> </table>	Mes	Casos	Costo	febrero	~15K	~\$300K	marzo	~19K	~\$450K	abril	~17K	~\$400K	mayo	~14K	~\$350K	junio	~19K	~\$450K	julio	~15K	~\$400K	agosto	~16K	~\$500K	septiembre	~13K	~\$300K	octubre	~14K	~\$300K	noviembre	~15K	~\$350K	diciembre	~22K	~\$500K	<p>Modismo: Line Chart Marcas: Puntos con marcas de conexión (líneas). Canales:</p> <ul style="list-style-type: none"> Color, para identificar cada atributo (casos y costo). Orden, ascendente por variable temporal (mes, día). Se utilizan etiquetas adicionales al ubicarse sobre cada categoría para ver la descripción completa de cada causa. <p>Decisiones de diseño:</p> <ul style="list-style-type: none"> Se usan los atributos de Casos y Costo, ya que son los principales motivadores de los planes de acción en el manejo de irregularidades. No se incluye el indicador MB en este mismo diagrama, ya que su escala es muy diferente a la de estos dos atributos y en la representación no es perceptible (se ubica junto con el eje horizontal de tiempo). Se usan etiquetas con el dato de pasajeros e indicador MB para consultar a demanda en cada punto del tiempo. 	T2, T7
Mes	Casos	Costo																																				
febrero	~15K	~\$300K																																				
marzo	~19K	~\$450K																																				
abril	~17K	~\$400K																																				
mayo	~14K	~\$350K																																				
junio	~19K	~\$450K																																				
julio	~15K	~\$400K																																				
agosto	~16K	~\$500K																																				
septiembre	~13K	~\$300K																																				
octubre	~14K	~\$300K																																				
noviembre	~15K	~\$350K																																				
diciembre	~22K	~\$500K																																				



Figura 27. Vista general de la solución para el análisis descriptivo

Consideraciones adicionales

Cabe resaltar que muchas de las tareas identificadas pueden ser resueltas con un mayor nivel de granularidad a través de la combinación de dos o más de estas gráficas. En el diseño general de la solución cada vista está coordinada con las demás, de tal manera que la interacción con cierta gráfica genera acciones de filtro o resaltado sobre las demás.

Es posible, por ejemplo, utilizar la primera gráfica para realizar un filtro por región, y realizar diversas tareas de análisis de tendencias, extremos, valores atípicos, etc. sobre las demás, sólo para esta región (Figura 28). También es posible realizarlas para un solo tipo de irregularidad en particular y sumar luego el filtro de región para analizar los saqueos en Norteamérica (Figura 29), por ejemplo.

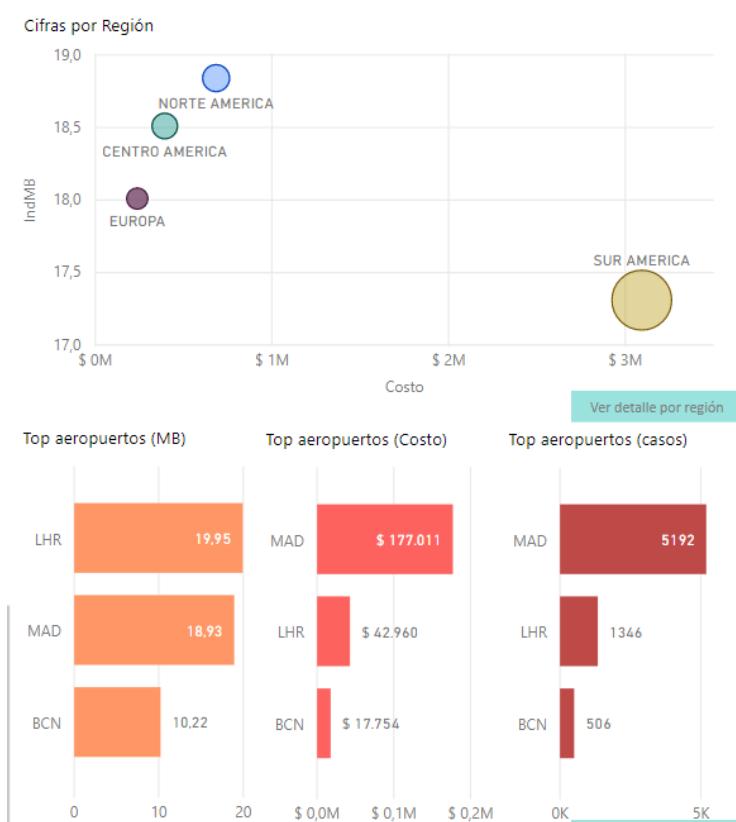


Figura 28. Coordinación de vistas (cifras por aeropuerto para la región Europa)

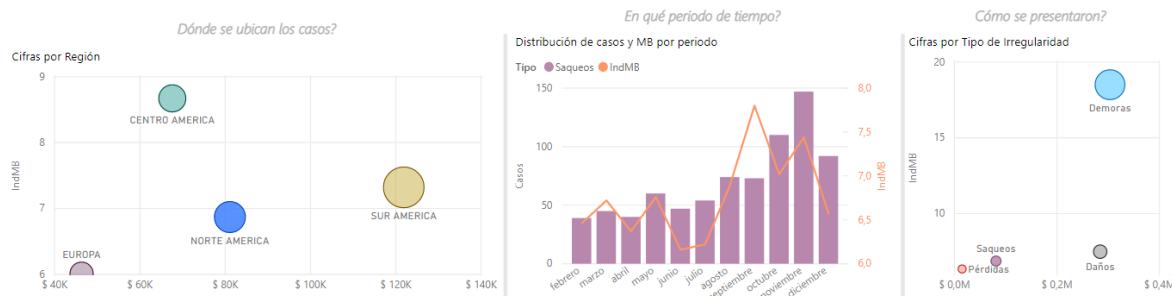


Figura 29. Coordinación de vistas (saqueos en Norteamérica)

Respecto a los diagramas que involucran tiempo, se disponen en alineación vertical para poder realizar la comparación de los diferentes indicadores en el mismo periodo (Figura 30), como compensación a la dificultad de utilizar todas las variables en un mismo diagrama por la diferencia en sus magnitudes y escalas.



Figura 30. Alineación de diagramas de tiempo

Por último, la disposición de la información presentada en la solución está asociada a la naturaleza de las preguntas de análisis que se pueden resolver (dónde, cuándo y cómo se presentan las irregularidades).

4.1.5. Evaluación

Durante la etapa de validación de la aproximación analítica utilizada, se identificaron algunos hallazgos relevantes, los cuales se describen a continuación.

4.1.5.1. *Hallazgos relevantes en datos*

Durante la exploración de los datos a través de la visualización se identificaron valores atípicos en el indicador MB. Considerando que este es un atributo generado a partir del número de irregularidades presentadas y el número de pasajeros transportados, se procedió a verificar este último. Se identificaron vuelos que, aunque tenían irregularidades asociadas, no contaban con el dato de los pasajeros transportados. También se identificaron vuelos con un número muy bajo de pasajeros (uno o dos). Esta situación eleva considerablemente el indicador. Para atender esta situación, se aplicó la técnica de imputación de datos sobre el atributo de pasajeros, usando como base otros vuelos con el mismo segmento del vuelo (origen-destino), tipo de aeronave y mes del recorrido.

4.1.5.2. *Hallazgos en el uso de modismos de visualización*

Una de las limitaciones encontradas en la etapa de diseño de la visualización fue la disponibilidad de modismos de visualización más sofisticados entre los objetos visuales de PowerBi. En algunos casos se encontraron propuestas de objetos visuales personalizados en Microsoft AppSource (Microsoft, 2020a), sin embargo factores como el costo de adquisición, el reducido tiempo de prueba de la versión *trial* o su restricción de uso solo para la versión de escritorio no permitieron contar con estos elementos.

Uno de estos casos se dio en el modismo de coordenadas paralelas, que se pretendía usar para presentar la distribución de las irregularidades, costos, indicador MB, pasajeros y maletas, por cada región, así como su correlación. Este modismo es una buena opción cuando se tienen varios atributos cuantitativos (Figura 31). El objeto visual se comportó muy bien en tiempo de diseño, no obstante al publicarlo en el servicio de PowerBi (cloud) para ser accedido por parte de los usuarios, el objeto no podía ser utilizado sin ser adquirido bajo un paquete mínimo de 10 usuarios por un periodo de 12 meses. Por esta razón, se eligió un *scatterplot* para manejar los tres atributos cuantitativos más relevantes del estudio (número de irregularidades, costos e indicador MB), dejando los demás para otros diagramas.

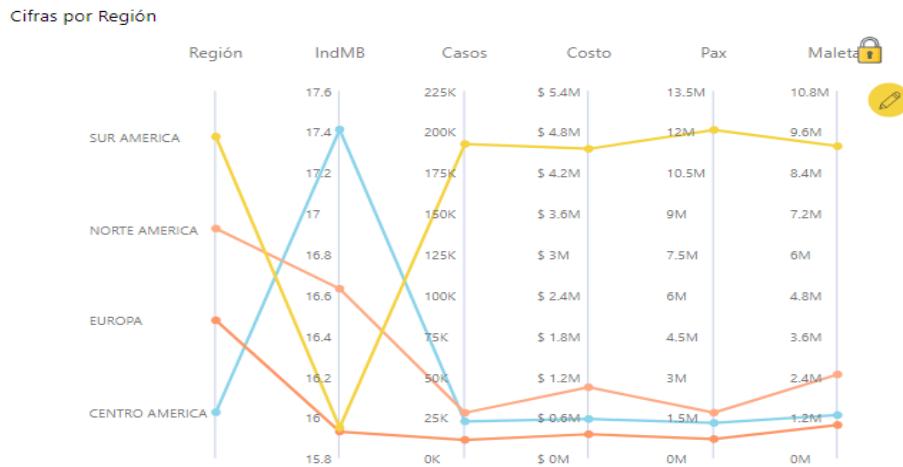


Figura 31. Coordenadas paralelas

Otro de los ajustes realizados en las diversas iteraciones de esta fase tuvo que ver con la selección del modismo de mapa para representar la distribución del número de irregularidades entre los aeropuertos responsables, así como el indicador MB en cada uno de ellos (Figura 32).

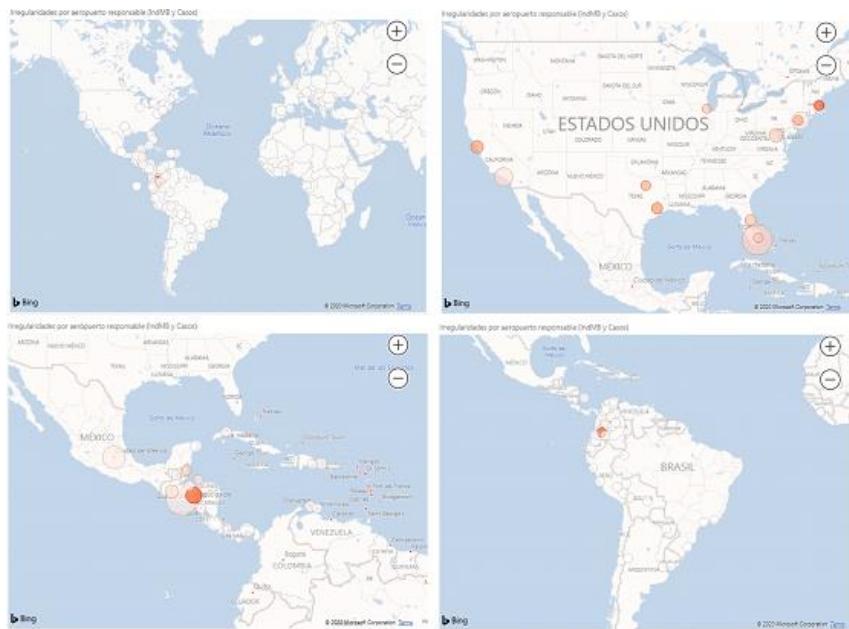


Figura 32. Mapa de aeropuertos responsables, irregularidades e indicador MB

El nivel de percepción de los aeropuertos con mayor número de casos (usando el tamaño de los puntos como canal) o con mayor indicador MB (usando el color como canal) no es satisfactorio en una vista del mapa completo. Es necesario realizar un acercamiento o filtro

por cada región. Aun así, en las regiones donde se da la mayor operación y, por tanto, donde se concentra el mayor número de irregularidades este problema se mantiene. Tal es el caso de Sur América. Regiones con baja operación y volumen de casos como Norte América o Europa tienen un mejor comportamiento. No obstante, por el tamaño considerable de espacio que este diagrama ocupa en la solución, dejando menos área para otras visualizaciones, se remplaza por otros diagramas que permiten alcanzar las tareas definidas (scatter plot para regiones y bar chart para aeropuertos, en ambos casos respecto del indicador MB y los costos como atributos principales, con coordinación de los dos diagramas).

4.1.6. Despliegue

El entorno de diseño y desarrollo local (descrito en la sección 4.1.4.1 Diseño de la herramienta) no es apto para la liberación de la solución a los usuarios en un entorno web. De otro lado, la publicación en el entorno *cloud* de PowerBi solo permite acceder a la cuenta que realiza el diseño. Por esta razón se utiliza la característica de Publicación en Web (Microsoft, 2020b). Es necesario tener en cuenta las características de este tipo de publicación: no usa autenticación, el contenido puede ser accedido por cualquier persona con la dirección URL, por tanto no se recomienda para entornos corporativos o cuando se maneja información confidencial. Para efectos de este proyecto la funcionalidad es suficiente, considerando que se emplean datos de prueba que no pueden ser considerados oficiales.

4.1.7. Retroalimentación

4.1.7.1. Pruebas funcionales

Las pruebas de usuario que brindan retroalimentación de la parte funcional de la solución se consolidan en la fase de análisis diagnóstico, por tanto serán descritas en la sección 4.2.7.

4.1.7.2. Pruebas de desempeño

En esta etapa se realizan mediciones de desempeño de la solución respecto al comportamiento de las interacciones, la coordinación entre visualizaciones y los tiempos de respuesta. Estas mediciones se realizan en entorno de desarrollo con la utilidad de *PowerBi*

Performance Analyzer, debido a que en el entorno de despliegue (PowerBi Service) no es posible obtener este tipo de mediciones para el nivel de la cuenta que se está utilizando.

Las mediciones se realizan realizando las tareas descritas en las pruebas de usuario y se analizan respecto de cada interacción (filtro, filtro combinado, *drill-down*, entre otros).

Las primeras interacciones con la solución arrojan unos tiempos elevados, no obstante el desempeño mejora con las siguientes y, en entorno de desarrollo, se identifica una mejora en los tiempos con un incremento proporcional en el uso de la memoria (especialmente en los servicios *CefSharp.BrowserSubprocess* y *Microsoft SQL Server Analysis Services*, de acuerdo con las estadísticas del sistema operativo). Las interacciones para las tareas en la *Vista de Resumen* generan un reporte de aquellos objetos visuales con los mayores tiempos en cada acción. Esto permite realizar algunas mejoras en el diseño de las interacciones, inactivando aquellas entre objetos que no hacen parte de las tareas (botones, etiquetas, imágenes, filtros) y modificando en algunos casos el resultado por el filtro entre objetos. La Tabla 12 muestra un ejemplo de las mediciones luego de los ajustes.

Tabla 12. Ejemplo de mediciones de tiempo en las interacciones

Acción	Medición																		
Acción de filtro simple	<table border="1"> <thead> <tr> <th data-bbox="768 1110 1209 1142">Name</th> <th data-bbox="1209 1110 1383 1142">Duration (ms)</th> </tr> </thead> <tbody> <tr> <td data-bbox="768 1142 1209 1174">⌚ Recording started (03/06/2020 10:23:47 a.m.)</td><td data-bbox="1209 1142 1383 1174">-</td></tr> <tr> <td data-bbox="768 1174 1209 1205">⌘ Cross-highlighted</td><td data-bbox="1209 1174 1383 1205">-</td></tr> <tr> <td data-bbox="768 1205 1209 1237">⊕ Indicador MB</td><td data-bbox="1209 1205 1383 1237">1211</td></tr> <tr> <td data-bbox="768 1237 1209 1269">⊕ Card</td><td data-bbox="1209 1237 1383 1269">1198</td></tr> <tr> <td data-bbox="768 1269 1209 1300">⊕ Top aeropuertos (casos)</td><td data-bbox="1209 1269 1383 1300">1194</td></tr> <tr> <td data-bbox="768 1300 1209 1332">⊕ Top aeropuertos (MB)</td><td data-bbox="1209 1300 1383 1332">1194</td></tr> <tr> <td data-bbox="768 1332 1209 1364">⊕ Top aeropuertos (Costo)</td><td data-bbox="1209 1332 1383 1364">1194</td></tr> <tr> <td data-bbox="768 1364 1209 1396">⊕ Distribución de casos y MB por periodo</td><td data-bbox="1209 1364 1383 1396">1194</td></tr> </tbody> </table>	Name	Duration (ms)	⌚ Recording started (03/06/2020 10:23:47 a.m.)	-	⌘ Cross-highlighted	-	⊕ Indicador MB	1211	⊕ Card	1198	⊕ Top aeropuertos (casos)	1194	⊕ Top aeropuertos (MB)	1194	⊕ Top aeropuertos (Costo)	1194	⊕ Distribución de casos y MB por periodo	1194
Name	Duration (ms)																		
⌚ Recording started (03/06/2020 10:23:47 a.m.)	-																		
⌘ Cross-highlighted	-																		
⊕ Indicador MB	1211																		
⊕ Card	1198																		
⊕ Top aeropuertos (casos)	1194																		
⊕ Top aeropuertos (MB)	1194																		
⊕ Top aeropuertos (Costo)	1194																		
⊕ Distribución de casos y MB por periodo	1194																		
Acción de filtro combinado	<table border="1"> <tbody> <tr> <td data-bbox="768 1400 1209 1431">⌘ Cross-highlighted</td> <td data-bbox="1209 1400 1383 1431">-</td> </tr> <tr> <td data-bbox="768 1431 1209 1463">⊕ Indicador MB</td> <td data-bbox="1209 1431 1383 1463">1097</td> </tr> <tr> <td data-bbox="768 1463 1209 1495">⊕ Distribución de casos y MB por periodo</td> <td data-bbox="1209 1463 1383 1495">1074</td> </tr> <tr> <td data-bbox="768 1495 1209 1526">⊕ Top aeropuertos (casos)</td> <td data-bbox="1209 1495 1383 1526">1070</td> </tr> <tr> <td data-bbox="768 1526 1209 1558">⊕ Top aeropuertos (Costo)</td> <td data-bbox="1209 1526 1383 1558">1070</td> </tr> <tr> <td data-bbox="768 1558 1209 1590">⊕ Top aeropuertos (MB)</td> <td data-bbox="1209 1558 1383 1590">1070</td> </tr> </tbody> </table>	⌘ Cross-highlighted	-	⊕ Indicador MB	1097	⊕ Distribución de casos y MB por periodo	1074	⊕ Top aeropuertos (casos)	1070	⊕ Top aeropuertos (Costo)	1070	⊕ Top aeropuertos (MB)	1070						
⌘ Cross-highlighted	-																		
⊕ Indicador MB	1097																		
⊕ Distribución de casos y MB por periodo	1074																		
⊕ Top aeropuertos (casos)	1070																		
⊕ Top aeropuertos (Costo)	1070																		
⊕ Top aeropuertos (MB)	1070																		
Acción de drill-down sobre línea de tiempo	<table border="1"> <tbody> <tr> <td data-bbox="768 1617 1209 1649">⌘ Cross-highlighted</td> <td data-bbox="1209 1617 1383 1649">-</td> </tr> <tr> <td data-bbox="768 1649 1209 1681">⊕ Indicador MB</td> <td data-bbox="1209 1649 1383 1681">1097</td> </tr> <tr> <td data-bbox="768 1681 1209 1712">⊕ Distribución de casos y MB por periodo</td> <td data-bbox="1209 1681 1383 1712">1074</td> </tr> <tr> <td data-bbox="768 1712 1209 1744">⊕ Top aeropuertos (casos)</td> <td data-bbox="1209 1712 1383 1744">1070</td> </tr> <tr> <td data-bbox="768 1744 1209 1776">⊕ Top aeropuertos (Costo)</td> <td data-bbox="1209 1744 1383 1776">1070</td> </tr> <tr> <td data-bbox="768 1776 1209 1807">⊕ Top aeropuertos (MB)</td> <td data-bbox="1209 1776 1383 1807">1070</td> </tr> </tbody> </table>	⌘ Cross-highlighted	-	⊕ Indicador MB	1097	⊕ Distribución de casos y MB por periodo	1074	⊕ Top aeropuertos (casos)	1070	⊕ Top aeropuertos (Costo)	1070	⊕ Top aeropuertos (MB)	1070						
⌘ Cross-highlighted	-																		
⊕ Indicador MB	1097																		
⊕ Distribución de casos y MB por periodo	1074																		
⊕ Top aeropuertos (casos)	1070																		
⊕ Top aeropuertos (Costo)	1070																		
⊕ Top aeropuertos (MB)	1070																		

Este ejercicio permite identificar los objetos con los tiempos más altos de respuesta en las consultas requeridas entre cada interacción. Aparecen de manera constante los tres diagramas de barras horizontales que representan el top de aeropuertos con los indicadores más altos (indicador MB, costo y número de casos). También el diagrama de gauge (medidor para el indicador MB). Con estos datos es posible tomar decisiones de diseño diferentes en siguientes iteraciones, como por ejemplo reducir el número de ítems en los diagramas (*top 3*, *top 5*) identificados, manteniendo la tarea para la cual fueron seleccionados (identificar los valores más altos).

4.2. Fase 2: Análisis Diagnóstico

4.2.1. Entendimiento del Negocio

En esta fase, cuyo principal objetivo es identificar elementos que puedan explicar las razones de posible causa (el por qué) respecto de los hechos ocurridos (las irregularidades), es necesario profundizar un poco más sobre el proceso de gestión de equipaje, descrito en la fase de análisis descriptivo. Durante las actividades de levantamiento de información se logran identificar diferentes características y datos relacionados con los equipajes, los cuales se presentan en la Figura 33.

Esta representación es de utilidad para abstraer los elementos, lugares, tiempos, pasos y personas que tienen relación con el equipaje durante su recorrido, así como para la identificación de necesidades de información y posibles fuentes de datos.

Durante esta etapa, se logran sintetizar algunas de las causas más conocidas de las irregularidades, de acuerdo con las entrevistas realizadas:

- Conexiones entre vuelos, ya sea de una misma aerolínea o de varias.
- Restricciones de peso o espacio de las aeronaves, que generan equipaje rezagado (no fue transportado en el mismo vuelo del pasajero).
- Revisiones de seguridad aeroportuaria, las cuales son más rigurosas en vuelos internacionales.
- Condiciones climáticas, que generan demoras en los vuelos.
- Gestión del equipaje registrado con mucha antelación y del equipaje rezagado (almacenamiento, enrutamiento).

- Errores en los sistemas de bandas de transporte (carruseles) y en el etiquetado de las maletas.
- Gestión del personal de operación terrestre y alta rotación de quienes ejercen funciones relacionadas con la carga-descarga manual de equipajes.

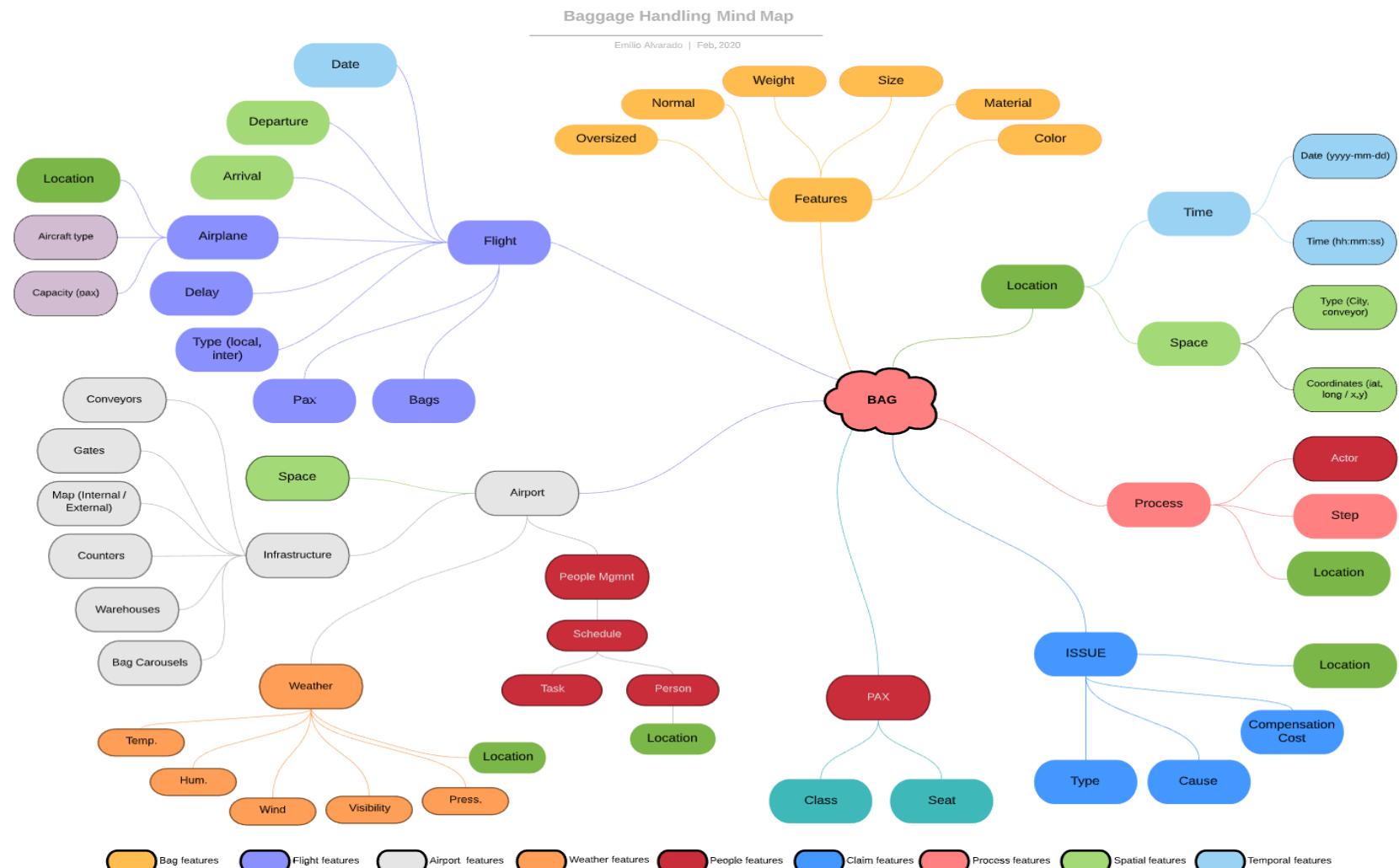


Figura 33. Mapa mental de información de contexto de los equipajes

Como paso siguiente se identificaron los sistemas de información que podrían alojar los datos relacionados con el proceso de equipajes y los demás elementos involucrados en la Figura 33. Estos sistemas se resumen en la Tabla 13. Se omiten los *Sistemas de manejo de equipajes* que fueron descritos en la sección 1.2.3.

Tabla 13. Sistemas involucrados en la gestión de equipajes

Sistema	Descripción
Sistema de Control de Salidas – Gestión del Cliente	Sistema que soporta los procesos y servicios a los pasajeros en relación con la salida de los vuelos (facturación, gestión de embarque, validación de tiquetes, autoservicio en check-in, gestión de políticas tarifarias, servicios complementarios). Un ejemplo de este tipo de sistemas es <i>Amadeus Customer Management</i> (Amadeus, 2019a).
Sistema de Control de Salidas – Gestión del Vuelo	Sistema que gestiona la carga de pasajeros y de cargamento en la aeronave. Incluye la optimización y distribución de la carga transportada, el cálculo del peso del combustible y la gestión del personal. Un ejemplo de este sistema es <i>Amadeus Flight Management</i> (Amadeus, 2019b).
Sistemas de Operación de Vuelos	Gestionan la información que soporta los procesos de operación y estado de vuelos, el manejo de la programación, los planes de mantenimiento, las desviaciones de los planes, entre otros. Un ejemplo de este tipo de solución es <i>Netline Ops</i> (Lufthansa-Systems, 2019).
Sistemas de Gestión del Relacionamiento con el Cliente	<i>Customer Relationship Management (CRM)</i> , administran la información del cliente y su relación con la organización, los productos que maneja, su categoría o nivel, los reclamos que ha registrado, las promociones que ha recibido, entre otros. Mantiene una vista de 360° del cliente para una administración centralizada. Un ejemplo de este tipo de sistema es <i>Microsoft Dynamics 365</i> (Microsoft, 2019a).
Sistema de Gestión de Irregularidades de Equipajes	Apoya el registro, seguimiento y gestión de las irregularidades presentadas por los clientes. Es un sistema generalmente desarrollado a la medida en las compañías de aviación. En otros casos se adquiere alguno de los productos de la suite de (SITA, 2019b).

4.2.2. Aproximación analítica

Se utilizaron tanto los métodos automatizados como la aproximación visual. Para la primera parte se aplicaron nuevamente los estadísticos básicos sobre los nuevos conjuntos de datos y particularmente se realizó un perfilamiento de los vuelos y las demoras presentadas respecto de su programación. Nuevamente, se aplicaron algunas correlaciones y se generaron histogramas y diagramas preliminares para apoyar el análisis descriptivo de los

nuevos conjuntos de datos (aquellos sobre los cuales se tuvo acceso a un conjunto de datos de prueba, más los que se obtuvieron de entidades y fuentes externas).

Para la segunda parte se extendió el tablero de control, agregando vistas adicionales para el análisis descriptivo y diagnóstico de aquellos conjuntos de datos que presentaban potencialmente una relación de causa con las irregularidades.

4.2.3. Procesamiento de Datos

En esta etapa se identificaron requerimientos de datos considerablemente mayores a aquellos de la fase descriptiva. Una de las mayores dificultades en esta etapa fue lograr el acceso a los sistemas mencionados en la sección 4.2.1 *Entendimiento del Negocio*, o al menos a un subconjunto de sus datos. Por lo anterior, se identificaron otras fuentes potenciales de información que pudieran apoyar la adquisición de datos para el estudio, especialmente las relacionadas con alguno de los factores que inciden en la ocurrencia de irregularidades. La Tabla 14 resume estas fuentes.

Tabla 14. Fuentes de información adicionales

Fuente	Descripción	Estado
Información de la flota de aeronaves	Codificación y clasificación de las aeronaves (dimensiones) y capacidad de pasajeros (asientos).	Codificación disponible en (AvCodes, 2019). Capacidad disponible en (SeatGuru, 2019) Dimensiones por clasificación, disponibles en (SKYbrary, 2020)
Información del personal de operación terrestre	Información de la planeación y ejecución de tareas del personal en el proceso de gestión de equipajes.	No disponible
Estadísticas e histórico de vuelos	Información de la programación, ejecución y estado de vuelos en un periodo y aeropuerto determinado.	Disponible en (FlightStats-Cirium, 2019)
Aviation-Edge	Información general de datos maestros como aerolíneas, aeropuertos, aeronaves, ciudades.	Disponible en (Aviation-Edge, 2019)
Reportes meteorológicos-METAR	Información del estado del tiempo en aeropuertos, específicamente a través del reporte METAR, altamente conocido en esta industria y estandarizado por la Organización de Aviación Civil	Disponible en (Aerocivil, 2020b)

Fuente	Descripción	Estado
	Internacional – ICAO (Wikipedia, 2019a). Reporte meteorológico de la Aerocivil.	
Estadísticas de las actividades aeronáuticas	Cifras consolidadas de la operación aeronáutica en Colombia, incluyendo histórico de vuelos realizados por ruta (origen-destino), tipo de aeronave, oferta y demanda de vuelos, boletines operacionales, quejas e información del servicio, tráfico en aeropuertos, entre otros.	Disponible en (Aerocivil, 2020a)
Estadísticas de las operaciones en el aeropuerto El Dorado	Información diaria consolidada de las operaciones, en reportes trimestrales, que incluyen tipo de vuelo (internacional, doméstico), cantidad de pasajeros que salen, aterrizajes y despegues.	Disponible en (OPAIN-S.A., 2020)
Información de la infraestructura Aeroportuaria	Información de la terminal 1 del aeropuerto El Dorado, puertas de abordaje, áreas de <i>make-up</i> , zona de equipaje entrante, zona equipaje de conexión, zona de equipaje saliente, sistema de transporte de equipaje.	Información parcial disponible en: Plan Operativo del Aeropuerto Internacional El Dorado (OPAIN-S.A., 2019) Mapa interno de El Dorado, zonas de reclamo de equipaje (Google, 2020b) Mapa externo de El Dorado, puertas de abordaje (Google, 2020a)

Al reunir el propósito identificado para esta fase de análisis diagnóstico (posibles causas de las irregularidades), las dimensiones en que se puede segregar el análisis (por periodo de tiempo, ubicación o causa RL asignada), las variables que influencian su ocurrencia y la disponibilidad de los datos para el análisis, se obtiene un panorama que consolida los elementos para abordar las siguientes etapas de la metodología, como se muestra en la Figura 34.

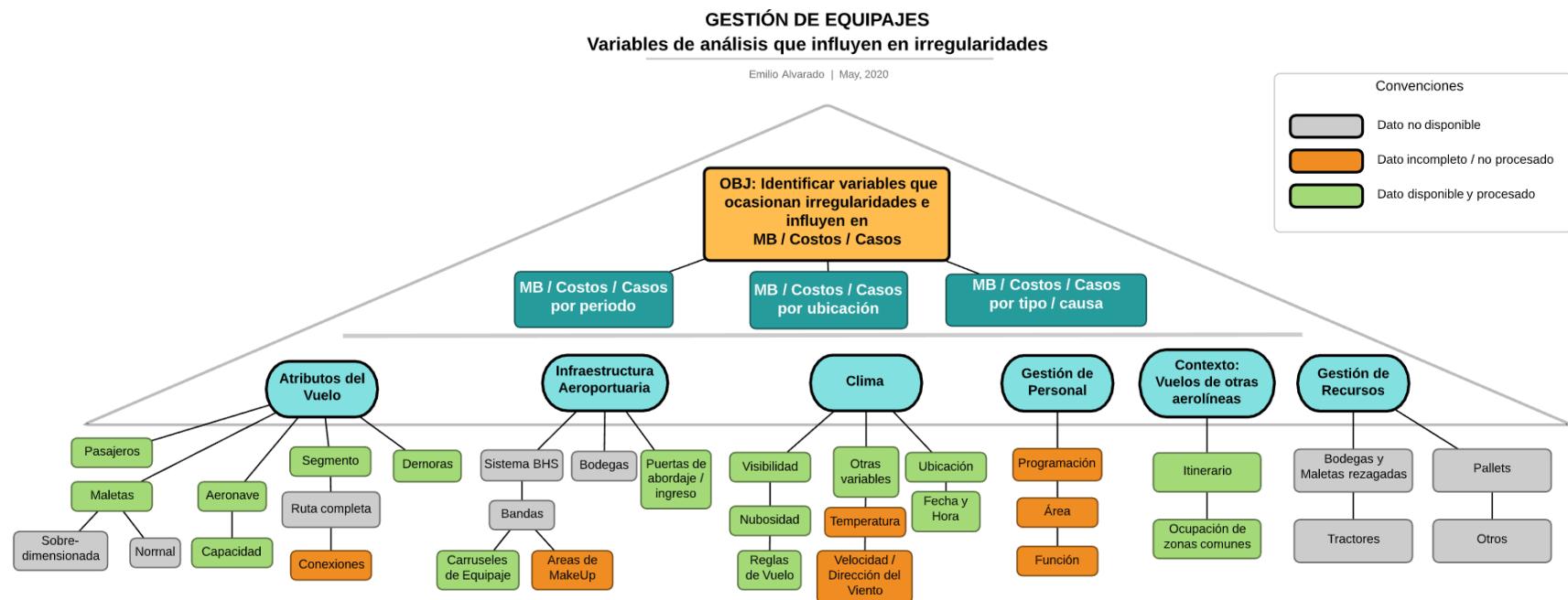


Figura 34. Variables de análisis y disponibilidad de datos

4.2.3.1. *Modelo conceptual de datos*

El modelo de datos en esta etapa se extendió para agregar algunas entidades adicionales que apoyan el análisis diagnóstico, desde las fuentes de información disponibles, como se muestra en la Figura 35. Se incluyen datos de vuelos y demoras presentadas, tipos y capacidad de aeronaves, estado del tiempo (Metar) y datos de contexto para el aeropuerto El Dorado, como las estadísticas de vuelos y pasajeros transportados y los vuelos de todas las aerolíneas para una fecha específica. También se realizó la generación de datos simulados para equipajes y su ubicación en el aeropuerto, con el fin de analizar estos datos en contexto para el mismo periodo de tiempo.

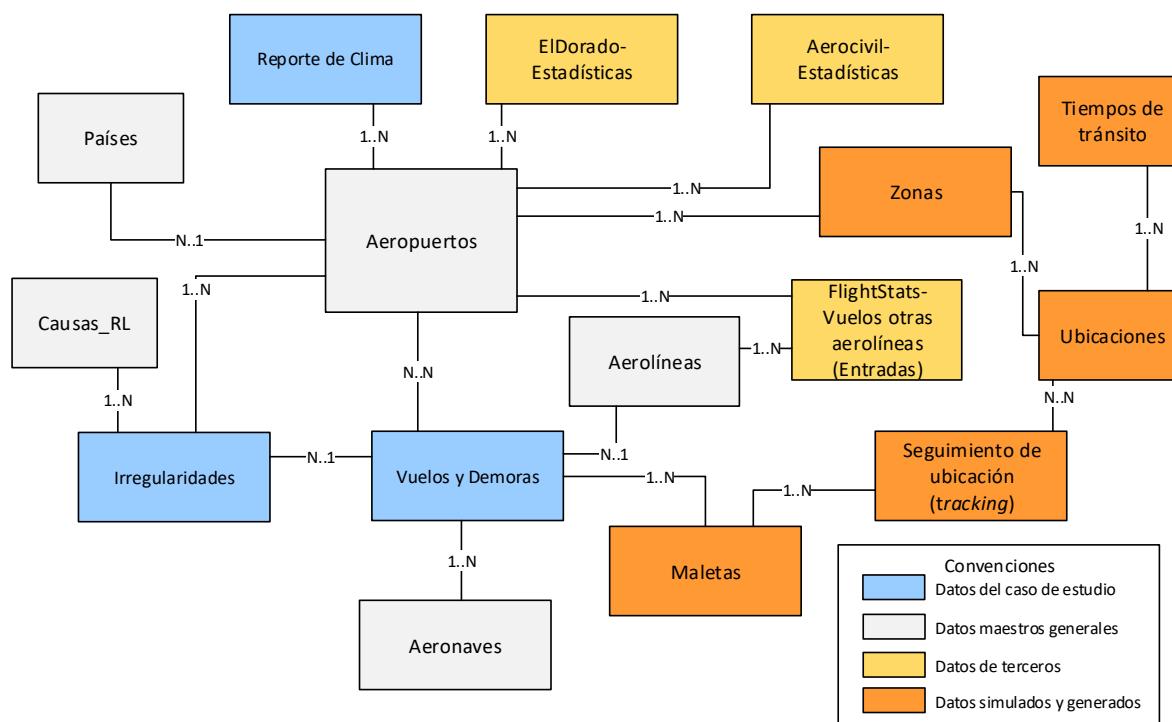


Figura 35. Modelo conceptual de datos en el análisis diagnóstico

Las relaciones están dadas por la siguiente lógica:

- A una irregularidad se asigna un solo aeropuerto responsable.
- En un aeropuerto se pueden presentar muchas irregularidades.
- Una irregularidad puede tener una sola causa (RL) y una causa puede ser asignada a muchas irregularidades.
- A una irregularidad se asigna un solo vuelo responsable y un vuelo puede ser causante de muchas irregularidades.

- Un aeropuerto opera muchos vuelos y un vuelo está relacionado con al menos un aeropuerto de origen y uno de destino.
- Un aeropuerto pertenece a un país y un país puede tener varios aeropuertos.
- Un aeropuerto puede tener muchos reportes de clima, y un reporte pertenece a un solo aeropuerto.
- Un vuelo puede presentar demoras, tanto en la salida como en la entrada.
- Un vuelo es comercializado por una aerolínea; una aerolínea comercializa muchos vuelos.
- Un vuelo puede tener muchas maletas; una maleta viaja en un solo vuelo².
- Un aeropuerto tiene muchas zonas; una zona pertenece a un solo aeropuerto.
- Una zona tiene muchas ubicaciones (puntos específicos); una ubicación pertenece a una sola zona.
- Una maleta puede estar en muchas ubicaciones; en una ubicación puede haber muchas maletas en un momento dado (*tracking*).
- Un aeropuerto tiene muchas estadísticas; las estadísticas pertenecen a un solo aeropuerto.

Las siguientes tablas resumen a nivel conceptual los conjuntos de datos obtenidos en esta etapa.

Tabla 15. Conjunto de datos de vuelos con demoras

Atributo	Tipo	Descripción
Número de vuelo	Numérico	Identificador del vuelo
Fecha de salida	Fecha-hora	Fecha programada del vuelo
Origen	Alfabético	Código del aeropuerto de origen
Destino	Alfabético	Código del aeropuerto de destino
Pasajeros	Numérico	Número de pasajeros transportados
Maletas	Numérico	Número de maletas transportadas
Tipo de Ruta	Alfabético	Indica si el vuelo es Doméstico o Internacional
Demora en la salida	Numérico	Minutos de demora respecto de la hora programada de salida
Demora en la llegada	Numérico	Minutos de demora respecto de la hora programada de llegada
Nivel de demora en salida	Numérico	0: vuelos con demora < a 15 minutos. 1: vuelos con demora entre 15 y 45 minutos. 2: vuelos con demora > a 45 minutos.

² Versión simplificada para este caso de estudio, se simulan datos para un solo día específico y no se incluye la variable de vuelos en conexiones para la generación y seguimiento de las maletas.

Atributo	Tipo	Descripción
Nivel de demora en llegada	Numérico	0: vuelos con demora < a 15 minutos. 1: vuelos con demora entre 15 y 45 minutos. 2: vuelos con demora > a 45 minutos.
Distancia	Numérico	Distancia de la ruta en Km
Duración	Numérico	Duración del vuelo en horas

Tabla 16. Conjunto de datos de aeropuertos

Atributo	Tipo	Descripción
Código IATA	Alfanumérico	Código del aeropuerto según la IATA
Código ICAO	Alfanumérico	Código del aeropuerto según la ICAO
GMT	Numérico	Indicador de conversión a tiempo universal (UTC) según la zona horaria del aeropuerto
Latitud	Numérico	Coordenada de latitud
Longitud	Numérico	Coordenada de longitud
Ciudad	Alfabético	Ciudad de ubicación del aeropuerto
País	Alfabético	País de ubicación del aeropuerto
Nombre	Alfabético	Nombre oficial del aeropuerto

Tabla 17. Conjunto de datos de clima

Atributo	Tipo	Descripción
Fecha de reporte	Fecha-hora	Fecha de lectura (en UTC)
Aeropuerto	Alfabético	Código ICAO del aeropuerto
Nubosidad	Alfabético	Tipo de nubes presentes según la lectura
Reglas de vuelo	Alfabético	Tipo de condiciones de vuelo según la visibilidad a la hora de la lectura
Temperatura	Numérico	Temperatura de la lectura, en °C
Visibilidad	Alfanumérico	Valor de la visibilidad al momento de la lectura
Velocidad el viento	Numérico	Valor de la velocidad del viento al momento de la lectura
Dirección del viento	Alfanumérico	Valor de la dirección del viento al momento de la lectura

Tabla 18. Conjunto de datos de aeronaves

Atributo	Tipo	Descripción
Código IATA	Alfanumérico	Código del tipo de aeronave según la IATA
Código ICAO	Alfanumérico	Código del tipo de aeronave según la ICAO
Capacidad	Numérico	Cantidad de asientos

Tabla 19. Conjunto de datos de aerolíneas

Atributo	Tipo	Descripción
Código IATA	Alfanumérico	Código de la aerolínea según la IATA
País	Alfanumérico	Código del país al que pertenece
Nombre	Alfabético	Nombre comercial de la aerolínea

Atributo	Tipo	Descripción
Estado	Alfabético	Vigencia de la aerolínea (si está operando, es un dato histórico, se encuentra suspendida, entre otros).

Tabla 20. Conjunto de datos de estadísticas de El Dorado

Atributo	Tipo	Descripción
Periodo reportado	Fecha	Fecha del reporte (año, mes, día)
Pax salida doméstico	Numérico	Cantidad de pasajeros transportados en la fecha del reporte para vuelos domésticos partiendo del aeropuerto.
Pax salida internacional	Numérico	Cantidad de pasajeros transportados en la fecha del reporte para destinos internacionales, partiendo del aeropuerto.
Aterrizajes	Numérico	Cantidad de vuelos ingresando a El Dorado en el día del reporte.
Despegues	Numérico	Cantidad de vuelos saliendo de El Dorado en el día del reporte.

Tabla 21. Conjunto de datos de estadísticas de Aerocivil

Atributo	Tipo	Descripción
Periodo reportado	Fecha	Fecha del reporte (año, mes)
Origen	Alfabético	Aeropuerto de salida
Destino	Alfabético	Aeropuerto de llegada
Tráfico	Alfabético	Tipo de ruta (nacional, internacional)
Pasajeros a bordo	Numérico	Cantidad de pasajeros transportados desde el origen
Pasajeros en tránsito	Numérico	Cantidad de pasajeros en tránsito.
Vuelos	Numérico	Cantidad de vuelos realizados en la ruta y periodo reportado
Sillas ofrecidas	Numérico	Cantidad de asientos ofertados por la aerolínea
Aerolínea	Alfanumérico	Código de la aerolínea que opera el vuelo
Tipo de equipo	Alfanumérico	Código del tipo de aeronave que realiza el vuelo
Carga ofrecida	Numérico	Volumen de carga (en kg.) ofertada por la aerolínea
Carga a bordo	Numérico	Volumen de carga (en kg.) transportada
Carga en tránsito	Numérico	Volumen de carga (en kg.) en tránsito

Tabla 22. Conjunto de datos de estadísticas de FlightStats

Atributo	Tipo	Descripción
Aerolínea comercial	Alfanumérico	Código de la aerolínea que comercializa el vuelo
Aerolínea que opera	Alfanumérico	Código de la aerolínea que opera el vuelo
Vuelo comercial	Numérico	Identificador del vuelo
Vuelo de operación	Numérico	Identificador del vuelo
Número de vuelo	Numérico	Identificador del vuelo

Atributo	Tipo	Descripción
Origen	Alfabético	Código del aeropuerto de origen
Destino	Alfabético	Código del aeropuerto de destino
Fecha de salida programada	Fecha-hora	Fecha programada del vuelo
Fecha de salida real	Fecha-hora	Fecha de salida real del vuelo
Estado	Alfabético	Estado final del vuelo
Aeronave	Alfanumérico	Código del tipo de aeronave
Puerta de salida	Alfanumérico	Código de la puerta de abordaje asignada para la salida
Puerta de llegada	Alfanumérico	Código de la puerta de abordaje asignada para la llegada
Carrusel	Alfanumérico	Código del carrusel de equipaje asignado al vuelo en su llegada

Además de los anteriores datos obtenidos, se generaron los siguientes conjuntos para el ejercicio de simulación tanto de ingreso como de salida de equipaje.

Tabla 23. Conjunto de datos de zonas

Atributo	Tipo	Descripción
Aeropuerto	Alfanumérico	Código IATA del aeropuerto
Código	Alfanumérico	Código de la zona
Nombre	Alfabético	Nombre de una zona específica del aeropuerto (check-in, escáner de seguridad, escáner secundario, áreas de make-up)
Cantidad de ubicaciones	Numérico	Cantidad de puntos o módulos por cada zona

Tabla 24. Conjunto de datos de ubicaciones

Atributo	Tipo	Descripción
Código de zona	Alfanumérico	Código de la zona a la que pertenece
Código de ubicación	Alfanumérico	Código de la ubicación
Latitud	Numérico	Coordenada de latitud
Longitud	Numérico	Coordenada de longitud

Tabla 25. Conjunto de datos de tiempos de tránsito

Atributo	Tipo	Descripción
Origen	Alfanumérico	Código de la ubicación de origen
Destino	Alfanumérico	Código de la ubicación de destino
Tiempo	Numérico	Tiempo de tránsito estimado

Tabla 26. Conjunto de datos de seguimiento de maletas

Atributo	Tipo	Descripción
Código de maleta	Alfanumérico	Código de la maleta (único por vuelo realizado)
Ubicación	Alfanumérico	Código de la ubicación

Atributo	Tipo	Descripción
Fecha	Fecha	Fecha de registro de la ubicación (año, mes, día)
Hora	Hora	Hora de registro de la ubicación (hora, minuto, segundo)

Tabla 27. Conjunto de datos de maletas

Atributo	Tipo	Descripción
Código de maleta	Alfanumérico	Código de la maleta (único por vuelo realizado)
Vuelo	Numérico	Código del vuelo en que se transporta
Fecha de salida	Fecha-hora	Fecha programada de salida
Origen	Alfabético	Aeropuerto de origen
Destino	Alfabético	Aeropuerto de destino
Tipo de ruta	Numérico	Indicador si es ruta doméstica o internacional

Por último, respecto a los datos relacionados con la infraestructura aeroportuaria, si bien no constituyen un conjunto de datos, los mapas referenciados en secciones anteriores aportan elementos de ubicación y cantidad de puntos dentro del aeropuerto que son de interés para el estudio. Tal es el caso de los carruseles de equipaje asignados a los vuelos (para el estudio de las maletas que ingresan), así como las puertas de abordaje. De otro lado, se tienen también las zonas de despacho de equipajes (*make-up areas*) para los vuelos que salen del aeropuerto El Dorado.

4.2.3.2. *Calidad y preparación de datos*

En los datos de vuelos se descartan algunos atributos como los comentarios de los pilotos, que dan una buena idea de los hechos que pudieron ocasionar las demoras, sin embargo el nivel de completitud no permite que sea un atributo a considerar para los análisis (99% de faltantes).

Aunque en los datos obtenidos no se tiene información de los vuelos en conexión, en el conjunto de datos de las irregularidades se cuenta con un atributo que indica la ruta completa que transitó el pasajero. A partir de este atributo se calcula el número de conexiones en los que participó el vuelo responsable de la irregularidad.

En el conjunto de datos de clima es necesaria la conversión en el tiempo, ya que los reportes manejan tiempo UTC y para hacerlos equivalentes a las horas de salida y llegada de los vuelos es necesario convertirlos a hora local. Con respecto a la visibilidad, hay ítems que están expresados en fracciones y millas, por lo cual es necesario estandarizar todos

los valores de lectura a números enteros expresados en km. En cuanto al aeropuerto donde se reporta el estado del tiempo, este se encuentra en codificación ICAO, es necesario homologarla a la codificación IATA presente en los demás conjuntos de datos.

Para efectos del análisis de clima se seleccionan únicamente los atributos de visibilidad, nubosidad y reglas de vuelo, considerando su estrecha relación con condiciones climáticas adversas que pueden ocasionar demoras en los vuelos.

Con respecto a los vuelos de otras aerolíneas que ingresan al aeropuerto El Dorado en una fecha específica (usando como fuente FlightStats), es necesario generar un atributo para identificar si el tipo de ruta es doméstica o internacional, para homologarla con los demás datos de vuelos y poder realizar análisis sobre este atributo. Adicionalmente, se estima el número de pasajeros transportados con base en la capacidad de asientos del tipo de aeronave asignada a cada vuelo. También se estima el número de maletas transportadas por cada vuelo, usando como base el conjunto de datos de prueba del caso de estudio, a partir de la relación entre pasajeros y maletas según el tipo de ruta del vuelo. Este mismo ejercicio se realiza para el conjunto de datos de estadísticas de Aerocivil, a partir del número de pasajeros y tipo de ruta se estima el número de maletas transportadas.

Sobre los demás conjuntos de datos nuevos se llevan a cabo tareas adicionales de estandarización de códigos, estandarización y asignación de valores nulos, identificación de duplicados e identificación y exclusión de atributos sin valor para el análisis o con niveles de calidad muy bajos.

4.2.3.3. Volumen de datos

Con respecto al volumen de datos, las cifras obtenidas a partir del procesamiento de datos se presentan en la Tabla 28. El conjunto de datos de mayor peso es el de irregularidades, tanto por el número de ítems como de atributos.

Tabla 28. Volumen de datos del análisis diagnóstico

Conjunto de datos	Ítems	Atributos	Tamaño
Irregularidades	390.203	28	84 MB
Vuelos y demoras	436.254	50	170 MB
Aeropuertos	9.000	13	1 MB
Países	252	5	88 KB
Causas_RL	48	5	2 KB
Aeronaves	22	4	1 KB
Aerolíneas	3962	14	1.1 MB

Conjunto de datos	Ítems	Atributos	Tamaño
Estadísticas de Aerocivil	25483	32	6.2 MB
Estadísticas El Dorado	365	8	58 KB
Reporte de Clima	1.188.209	14	136 MB
Vuelos de otras aerolíneas (<i>FlightStats</i>)	577	58	266 KB
Zonas	4	5	1 KB
Ubicaciones	20	5	1 KB
Maletas	23.241	8	1.8 MB
Ubicación de maletas (tracking)	92.964	4	4.7 MB
Tiempos de tránsito	87	3	2 KB

4.2.3.4. Entendimiento de los datos

En esta etapa se involucran algunos datos que es necesario comprender en contexto para el objeto del análisis. Este es el caso de las variables del clima, cuyos valores corresponden a codificaciones estandarizadas en la industria. No se pretende ser exhaustivo en la interpretación de estas variables, sino incorporar los elementos más relevantes que permitan determinar si un reporte de clima puede interpretarse bajo condiciones adversas que afecten la operación de los vuelos y, en relación con los análisis sobre los datos, si bajo estas condiciones efectivamente los vuelos presentaron demoras e irregularidades.

En términos de los valores en el atributo visibilidad y en las reglas de vuelo, la interpretación a utilizar se describe en la Tabla 29.

Tabla 29. Interpretación de los valores para reglas de vuelo y visibilidad

Acrónimo	Significado	Visibilidad
VFR	<i>Visual Flight Rules</i>	Mayor a 5 millas (8 km)
MVFR	<i>Marginal Visual Flight Rules</i>	Entre 3 y 5 millas (4.8 a 8 km)
IFR	<i>Instrumental Flight Rules</i>	Entre 1 y menos de 3 millas (1.6 a menos de 4.8 km)
LIFR	<i>Low Instrumental Flight Rules</i>	Menor a 1 milla (1.6 km)

Respecto de la nubosidad, se entiende como la fracción del cielo cubierto con nubes en un lugar en particular, expresada en octas u octavos de la bóveda celeste. Al evaluar el número de partes que están cubiertas con nubes, se estima el rango de visibilidad del observador. Así, entre más partes cubiertas con nubes menor es la visibilidad. La Tabla 30 resume los valores para este atributo y su interpretación (Wikipedia, 2019a).

Tabla 30. Interpretación de los valores para el atributo nubosidad

Acrónimo	Significado	Símbolo
SKC	"No cloud/Sky Clear" se usa en todo el mundo, pero en América del Norte se usa para indicar un informe generado por personas.	○ SKC

Acrónimo	Significado	Símbolo
NCD	"Nil Cloud Detected" la estación METAR automatizada no ha detectado ninguna nube, ya sea por falta de ella o por un error en los sensores.	
CLR	CLEAR. No hay nubes por debajo de los 12.000 pies (3.700 mts)	□ CLR
NSC	"No (nil) Significant Cloud", ninguna nube por debajo de los 5.000 pies (1.500 mts).	
FEW	"Few" = 1–2 octas (pocas)	① FEW
SCT	"Scattered" = 3–4 octas (disperso)	○ SCT
BKN	"Broken" = 5–7 octas (roto)	● BKN
OVC	"Overcast" = 8 octas, es decir, cobertura total (nublado)	● OVC
VV	Vertical Visibility. Las nubes no se pueden ver debido a la niebla o las fuertes precipitaciones.	

4.2.4. Modelamiento y Visualización

Esta sección presenta los temas relacionados con el diseño de la solución, los análisis automatizados sobre los datos y el diseño de la herramienta de visualización.

4.2.4.1. Diseño de la herramienta

El diseño de alto nivel se mantiene como fue presentado en la fase de análisis descriptivo (Figura 24). Para la implementación de los elementos de la vista lógica en esta fase fue necesario incrementar las capacidades de procesamiento debido al volumen de datos y a la exigencia en tiempo de ejecución tanto de los scripts de generación de datos como el diseño de la visualización. El entorno local fue migrado a un entorno de nube, con servicios de almacenamiento en Google Drive y servicios PaaS en lenguaje Python con Google Colab. Para la parte de visualización se utilizó un modelo IaaS (una máquina virtual en los servidores de la universidad) en tiempo de diseño, y el servicio público de PowerBi para el despliegue. La Figura 36 muestra la vista física de esta implementación.

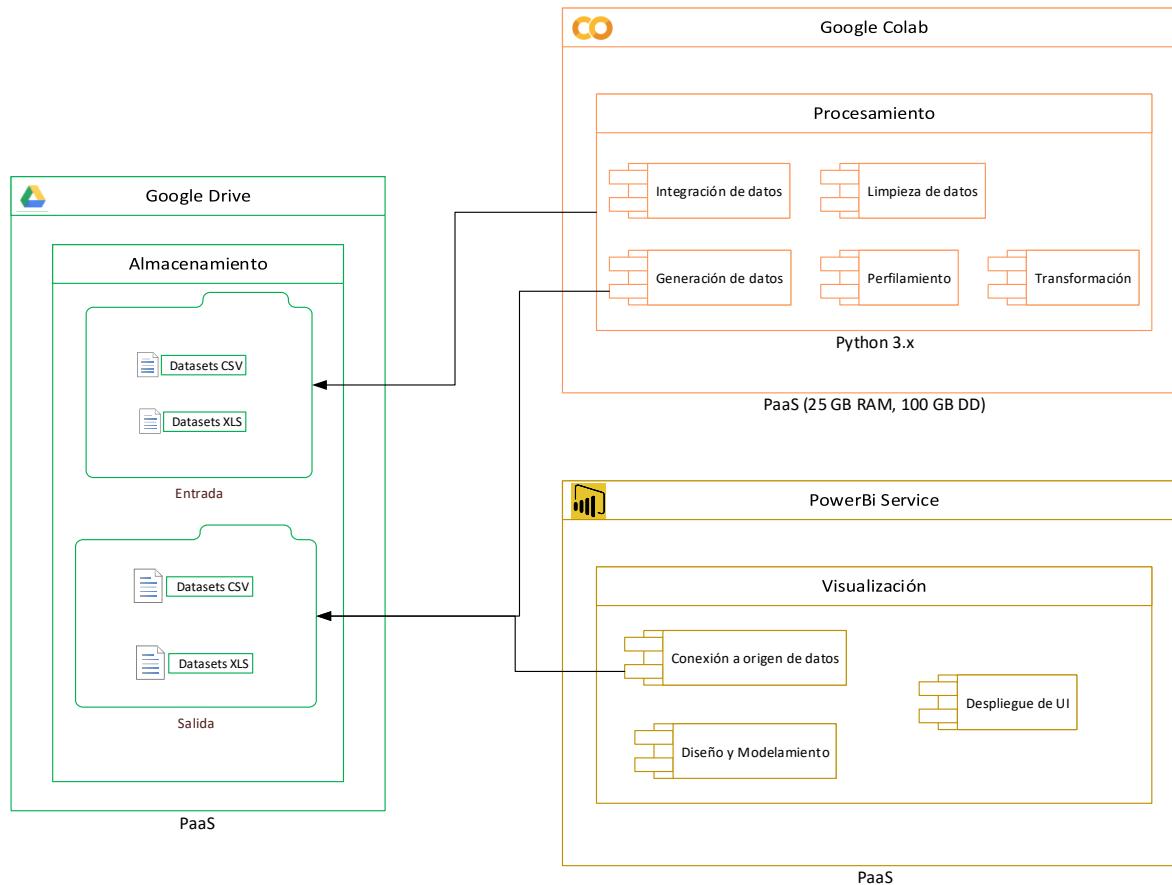


Figura 36. Vista física de la solución – análisis diagnóstico

4.2.4.2. Análisis asistido

Se realizó un análisis descriptivo exploratorio sobre los nuevos conjuntos de datos, especialmente los datos de demoras en los vuelos, los datos de estado del tiempo y los datos de estadísticas de terceros. También se realizan algunas correlaciones básicas sobre los atributos numéricos de los vuelos analizados con ubicación en el aeropuerto El Dorado, como se muestra en la Figura 37.

	DURACION	DISTANCIA	DEPARTURE_DELAY	BAGS	PAX	DEMORA	PERDIDA	DANO	SAQUEO	TOTAL_CASOS	MB
DURACION	1.000000	0.993997	0.069406	0.710558	0.431488	0.190968	0.042095	0.156521	0.141089	0.221013	0.104557
DISTANCIA	0.993997	1.000000	0.068044	0.717684	0.444867	0.184176	0.041304	0.160040	0.139831	0.214949	0.097102
DEPARTURE_DELAY	0.069406	0.068044	1.000000	0.061640	0.069034	0.050589	0.004926	0.023407	0.016367	0.053824	0.032057
BAGS	0.710558	0.717684	0.061640	1.000000	0.727541	0.197599	0.045223	0.165094	0.148132	0.229400	0.070535
PAX	0.431488	0.444867	0.069034	0.727541	1.000000	0.112994	0.025360	0.137471	0.090992	0.138713	-0.026536
DEMORA	0.190968	0.184176	0.050589	0.197599	0.112994	1.000000	0.025709	0.089715	0.087458	0.982985	0.802776
PERDIDA	0.042095	0.041304	0.004926	0.045223	0.025360	0.025709	1.000000	0.011158	0.021046	0.046431	0.031471
DANO	0.156521	0.160040	0.023407	0.165094	0.137471	0.089715	0.011158	1.000000	0.029820	0.252431	0.167544
SAQUEO	0.141089	0.139831	0.016367	0.148132	0.090992	0.087458	0.021046	0.029820	1.000000	0.168116	0.108968
TOTAL_CASOS	0.221013	0.214949	0.053824	0.229400	0.138713	0.982985	0.046431	0.252431	0.168116	1.000000	0.808003
MB	0.104557	0.097102	0.032057	0.070535	-0.026536	0.802776	0.031471	0.167544	0.108968	0.808003	1.000000

Figura 37. Correlación entre atributos numéricos en análisis diagnóstico

Es posible identificar algunas correspondencias esperadas, como la que se presenta entre la duración y la distancia de los vuelos, el número de casos y el indicador MB, las irregularidades de tipo demoras y el indicador MB (al ser este tipo de irregularidad la que más predomina en todos los casos), el número de pasajeros y el número de maletas, y estos dos acorde a la duración y distancia del vuelo.

4.2.4.3. *Diseño de la visualización*

ABSTRACCIÓN DE LOS DATOS

El tipo de conjunto de datos que predomina sigue siendo el de tablas, con disponibilidad estática.

Los tipos de atributos identificados en este punto para los conjuntos de datos nuevos o extendidos se describen a continuación:

Tabla 31. Tipos de atributos de los vuelos y demoras

Atributo	Tipo	Dirección / Ordenamiento	Jerarquía
Número de vuelo	Cuantitativo	Secuencial	NA
Fecha de salida	Ordinal (usado como fecha) Cuantitativo (usado como día, mes y año por separado)	Cíclico	Si (Año / Mes / Día)
Origen	Categórico	NA	No
Destino	Categórico	NA	No
Pasajeros	Cuantitativo	Secuencial	NA
Maletas	Cuantitativo	Secuencial	NA
Tipo de Ruta	Categórico	NA	No
Demora en la salida	Cuantitativo	Divergente	No
Demora en la llegada	Cuantitativo	Divergente	No
Nivel de demora en salida	Categórico	Secuencial	No
Nivel de demora en llegada	Categórico	Secuencial	No
Distancia	Cuantitativo	Secuencial	No
Duración	Cuantitativo	Secuencial	No

Tabla 32. Tipos de atributos de aeropuertos

Atributo	Tipo	Dirección / Ordenamiento	Jerarquía
Código IATA	Categórico	NA	No

Atributo	Tipo	Dirección / Ordenamiento	Jerarquía
Código ICAO	Categórico	NA	No
GMT	Cuantitativo	Divergente	No
Latitud	Cuantitativo	Divergente	No
Longitud	Cuantitativo	Divergente	No
Ciudad	Categórico	NA	Si
País	Categórico	NA	Si
Nombre	Categórico	NA	Si

Tabla 33. Tipos de atributos del clima

Atributo	Tipo	Dirección / Ordenamiento	Jerarquía
Fecha de reporte	Ordinal (usado como fecha) Cuantitativo (usado como día, mes y año por separado)	Cíclico	Si (Año / Mes / Día)
Aeropuerto	Categórico	NA	No
Nubosidad	Categórico	NA	No
Reglas de vuelo	Categórico	NA	No
Temperatura	Cuantitativo	Divergente	No
Visibilidad	Cuantitativo	Secuencial	No
Velocidad el viento	Cuantitativo	Secuencial	No
Dirección del viento	Cuantitativo	Divergente	No

Tabla 34. Tipos de atributos de las aeronaves

Atributo	Tipo	Dirección / Ordenamiento	Jerarquía
Código IATA	Categórico	NA	No
Código ICAO	Categórico	NA	No
Capacidad	Cuantitativo	Secuencial	No

Tabla 35. Tipos de atributos de aerolíneas

Atributo	Tipo	Dirección / Ordenamiento	Jerarquía
Código IATA	Categórico	NA	No
País	Categórico	NA	No
Nombre	Categórico	NA	No
Estado	Categórico	NA	No

Tabla 36. Tipos de atributos para las estadísticas de El Dorado

Atributo	Tipo	Dirección / Ordenamiento	Jerarquía
Periodo reportado	Ordinal (usado como fecha) Cuantitativo (usado como día, mes y año por separado)	Cíclico	Si (Año / Mes / Día)
Pax salida domesticos	Cuantitativo	Secuencial	No
Pax salida internacional	Cuantitativo	Secuencial	No
Aterrizajes	Cuantitativo	Secuencial	No
Despegues	Cuantitativo	Secuencial	No

Tabla 37. Tipos de atributos para las estadísticas de Aerocivil

Atributo	Tipo	Dirección / Ordenamiento	Jerarquía
Periodo reportado	Ordinal (usado como fecha) Cuantitativo (usado como día, mes y año por separado)	Cíclico	Si (Año / Mes / Día)
Origen	Categórico	NA	No
Destino	Categórico	NA	No
Tráfico	Categórico	NA	No
Pasajeros a bordo	Cuantitativo	Secuencial	No
Pasajeros en tránsito	Cuantitativo	Secuencial	No
Vuelos	Cuantitativo	Secuencial	No
Sillas ofrecidas	Cuantitativo	Secuencial	No
Aerolínea	Categórico	NA	No
Tipo de equipo	Categórico	NA	No
Carga ofrecida	Cuantitativo	Secuencial	No
Carga a bordo	Cuantitativo	Secuencial	No
Carga en tránsito	Cuantitativo	Secuencial	No

Tabla 38. Tipos de atributos para las estadísticas de FlightStats

Atributo	Tipo	Dirección / Ordenamiento	Jerarquía
Aerolínea comercial	Categórico	NA	No
Aerolínea que opera	Categórico	NA	No
Vuelo comercial	Categórico	Si	No
Vuelo de operación	Categórico	Si	No
Origen	Categórico	NA	No
Destino	Categórico	NA	No
Fecha de salida programada	Ordinal (usado como fecha) Cuantitativo (usado como día, mes y año por separado)	Cíclico	Si (Año / Mes / Día)

Atributo	Tipo	Dirección / Ordenamiento	Jerarquía
Fecha de salida real	Ordinal (usado como fecha) Cuantitativo (usado como día, mes y año por separado)	Cíclico	Si (Año / Mes / Día)
Estado	Categórico	NA	No
Aeronave	Categórico	NA	No
Puerta de salida	Categórico	NA	No
Puerta de llegada	Categórico	NA	No
Carrusel	Categórico	Si	No

Tabla 39. Tipos de atributos para las Zonas

Atributo	Tipo	Dirección / Ordenamiento	Jerarquía
Aeropuerto	Categórico	NA	No
Código	Categórico	NA	No
Nombre	Categórico	NA	No
Cantidad de ubicaciones	Cuantitativo	Secuencial	No

Tabla 40. Tipos de atributos para las Ubicaciones

Atributo	Tipo	Dirección / Ordenamiento	Jerarquía
Código de zona	Categórico	NA	No
Código de ubicación	Categórico	NA	No
Latitud	Cuantitativo	Divergente	No
Longitud	Cuantitativo	Divergente	No

Tabla 41. Tipos de atributos para el seguimiento de maletas

Atributo	Tipo	Dirección / Ordenamiento	Jerarquía
Código de maleta	Categórico	Si	No
Ubicación	Categórico	NA	No
Fecha	Ordinal (usado como fecha) Cuantitativo (usado como día, mes y año por separado)	Cíclico	Si (Año / Mes / Día)
Hora	Ordinal (usado como tiempo) Cuantitativo (usado como hora, minuto y segundo por separado)	Cíclico	Si (Hora/ Minuto/ Segundo)

Tabla 42. Tipos de atributos para las maletas

Atributo	Tipo	Dirección / Ordenamiento	Jerarquía
Código de maleta	Categórico	Si	No
Vuelo	Categórico	Si	No
Fecha de salida	Ordinal (usado como fecha) Cuantitativo (usado como día, mes y año por separado)	Cíclico	Si (Año / Mes / Día)
Origen	Categórico	No	No
Destino	Categórico	No	No
Tipo de ruta	Categórico	No	No

ABSTRACCIÓN DE LAS TAREAS

Los perfiles de usuario objetivo identificados anteriormente como **Coordinador de la Oficina Central de Rastreo de Equipajes** (localizado en el aeropuerto) y **Gerente del Departamento de Gestión de Equipajes** (de parte de las aerolíneas) permanecen vigentes.

Respecto de las tareas identificadas para estos actores, descritas en términos de **acciones** y **objetivos**, a continuación se enumeran las que corresponden a esta fase de análisis diagnóstico:

- T8. Localizar las **características** de los vuelos que inciden en el indicador MB.
- T9. Explorar la **correlación** entre las variables de los vuelos, las variables del clima y las irregularidades presentadas (en términos del indicador MB y el número de casos).
- T10. Presentar la **correlación** entre los vuelos de otras aerolíneas y los propios, y su relación con las irregularidades presentadas.
- T11. Localizar **valores atípicos** y **características** de los vuelos que ingresan al aeropuerto El Dorado, y su relación con las irregularidades presentadas.
- T12. Localizar **valores atípicos** y **características** de los vuelos que salen del aeropuerto El Dorado, y su relación con las irregularidades presentadas.

DECISIONES DE DISEÑO

En esta fase se incluyeron vistas adicionales para agrupar las tareas identificadas. En cada vista se incluyen algunos gráficos de la fase de análisis descriptivo para contar con un punto de partida en la ubicación del lugar, periodo de tiempo y tipo de irregularidad en que se

desea enfocar el análisis. Luego se agregan las nuevas visualizaciones que responden a las tareas de la fase diagnóstica, usando las anteriores como filtros e incluyendo coordinación entre ellas. Esta es una de las decisiones generales adoptadas con el fin de contar en cada vista con la posibilidad de aplicar el *visual information-seeking mantra* descrito en (Shneiderman, 2000), de la siguiente manera:

- *Overview*: se incluyen diagramas que dan un estado general de la nueva variable a analizar (demoras en vuelos, estado del tiempo, tendencia de los indicadores en el año, por ejemplo). También se presentan etiquetas con las cifras más relevantes (indicador MB, costo, casos, pasajeros y maletas) en la parte superior, las cuales se actualizan en toda interacción. Solamente se maneja un medidor de *Gauge* para el indicador MB por ser el atributo que se considera el más relevante y sobre el cual se tiene indicio del rango objetivo de valores que puede adoptar.
- *Zoom & Filter*: se agregan filtros sobre atributos específicos en la zona superior, y se mantiene la coordinación de vistas entre todos los objetos (salvo contadas excepciones).
- *Details on Demand*: Ciertos objetos (como los diagramas de barras o líneas para tendencias en el año, los *Treecharts* o las matrices) manejan jerarquías en las cuales es posible ampliar el nivel de detalle realizando acciones de *drill-down* entre los niveles (por ejemplo, de mes a día, de tipo de causa a código de causa, de segmento volado a vuelo específico). Adicionalmente todos los diagramas presentan etiquetas para obtener información adicional en cada ítem de datos.

A continuación se describen las consideraciones de diseño por cada vista y tarea identificada, incluyendo los modismos de visualización, canales, marcas, decisiones y observaciones.

La Figura 38 presenta la disposición general de la vista de Vuelos y Demoras, y la Tabla 43 describe las decisiones de diseño en esta vista.

La Figura 39 presenta la disposición general de la vista de Condiciones Climáticas, y la Tabla 44 describe las decisiones de diseño en esta vista.

La Figura 40 presenta la disposición general de la Vista de Contexto, y la Tabla 45 describe las decisiones de diseño en esta vista.

La Figura 41 presenta la disposición general de la Vista de Contexto – Vuelos que ingresan a El Dorado (*Entradas-BOG*), y la Tabla 46 describe las decisiones de diseño en esta vista.

La Figura 42 presenta la disposición general de la Vista de Contexto – Vuelos saliendo de El Dorado (*Salidas-BOG*), y la Tabla 47 describe las decisiones de diseño en esta vista.

Po último, la Figura 43 presenta la disposición general de la Guía de Usuario.

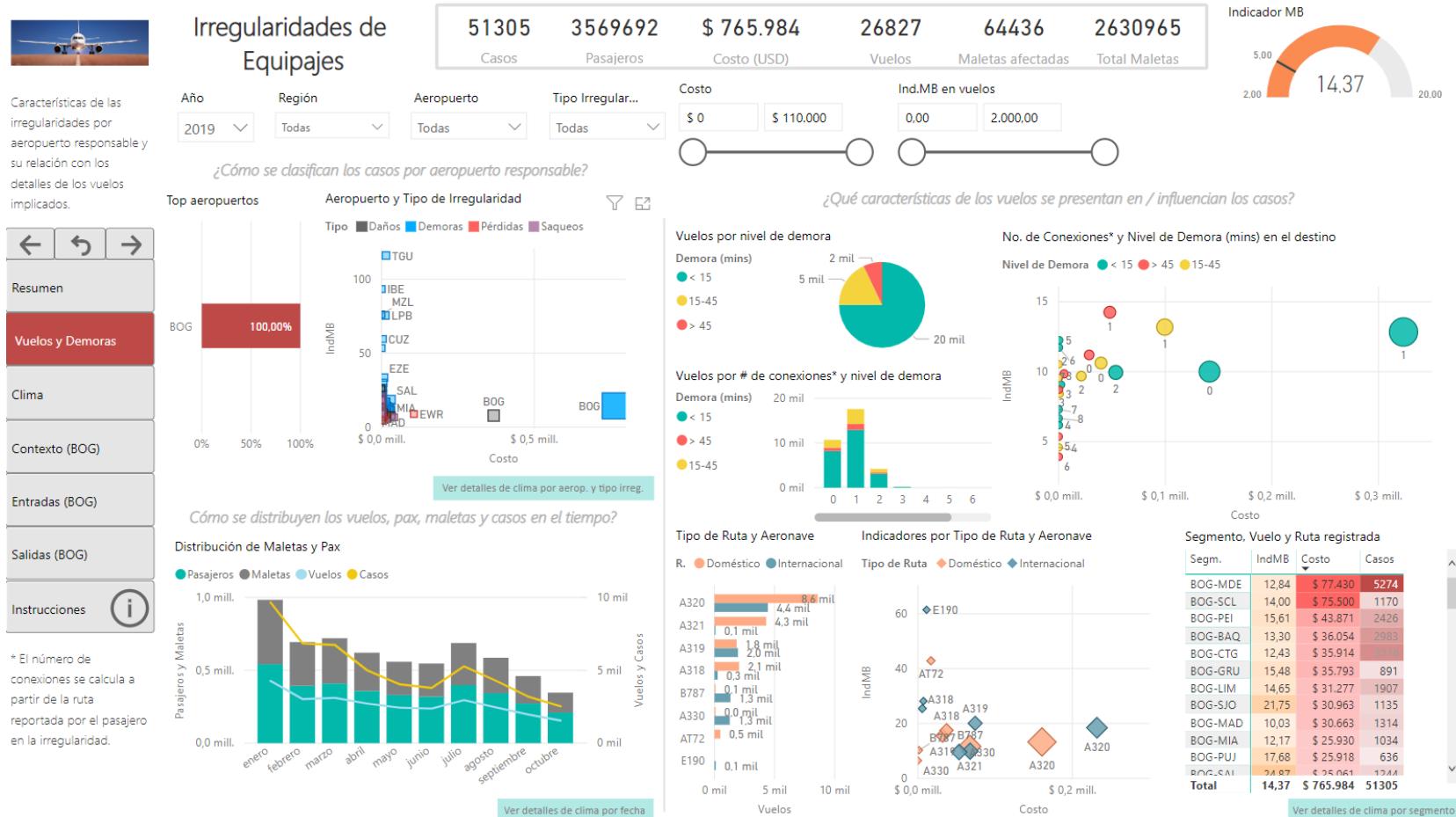


Figura 38. Vista de Vuelos y Demoras

Tabla 43. Decisiones de diseño para la vista de vuelos y demoras

Representación	Codificación	Tarea																			
<p>¿Cómo se clasifican los casos por aeropuerto responsable?</p> <p>Top aeropuertos</p> <table border="1"> <thead> <tr> <th>Aeropuerto</th> <th>Porcentaje</th> </tr> </thead> <tbody> <tr><td>BOG</td><td>64.59%</td></tr> <tr><td>MIA</td><td>3.66%</td></tr> <tr><td>MAD</td><td>3.46%</td></tr> <tr><td>SAL</td><td>3.08%</td></tr> <tr><td>LIM</td><td>2.46%</td></tr> <tr><td>MEX</td><td>1.87%</td></tr> <tr><td>LAX</td><td>1.68%</td></tr> <tr><td>JFK</td><td>1.54%</td></tr> <tr><td>SCL</td><td>1.39%</td></tr> </tbody> </table> <p>Aeropuerto y Tipo de Irregularidad</p> <p>IndMB</p> <p>Tipo</p> <ul style="list-style-type: none"> Datos Demoras Pérdidas Saqueos <p>Costo</p> <p>Ver detalles de clima por aerop. y tipo irreg.</p> <p>Distribución de Maletas y Pax</p> <p>Pasajeros y Maletas</p> <p>Vuelos y Casos</p> <p>enero febrero marzo abril mayo junio julio agosto septiembre octubre</p> <p>IndMB</p> <p>Modismos: Pie Chart, Vertical Stacked Bar Chart, Scatterplot</p> <p>Marcas: Áreas, líneas, puntos (respectivamente)</p> <p>Canales (barchart):</p> <ul style="list-style-type: none"> Color, para identificar el nivel de demora de los vuelos (todos los diagramas). Longitud, para identificar el número de vuelos en cada barra. Orden, de acuerdo con el número de conexiones. <p>Canales (Scatterplot):</p> <ul style="list-style-type: none"> Posición en X para el costo. Posición en Y para el indicador MB Tamaño para el número de casos <p>Decisiones de diseño:</p> <ul style="list-style-type: none"> Estas visualizaciones adoptan tareas de la fase descriptiva, con el fin de ubicar el punto de partida para el análisis diagnóstico. Los puntos de partida para el análisis pueden ser: el aeropuerto responsable (por número de casos), la distribución de aeropuertos por indicador MB y costos o la tendencia en el año de los casos y los volúmenes de los vuelos, pasajeros y maletas. 	Aeropuerto	Porcentaje	BOG	64.59%	MIA	3.66%	MAD	3.46%	SAL	3.08%	LIM	2.46%	MEX	1.87%	LAX	1.68%	JFK	1.54%	SCL	1.39%	T1, T2, T5
Aeropuerto	Porcentaje																				
BOG	64.59%																				
MIA	3.66%																				
MAD	3.46%																				
SAL	3.08%																				
LIM	2.46%																				
MEX	1.87%																				
LAX	1.68%																				
JFK	1.54%																				
SCL	1.39%																				
<p>¿Qué características de los vuelos se presentan en / influyen los casos?</p> <p>Vuelos por nivel de demora</p> <p>Demora (mins)</p> <ul style="list-style-type: none"> < 15 15-45 > 45 <p>Vuelos por # de conexiones* y nivel de demora</p> <p>Demora (mins)</p> <ul style="list-style-type: none"> < 15 > 45 15-45 <p>No. de Conexiones* y Nivel de Demora (mins) en el destino</p> <p>Nivel de Demora</p> <ul style="list-style-type: none"> < 15 > 45 15-45 <p>IndMB</p> <p>Decisiones de diseño:</p> <ul style="list-style-type: none"> Este conjunto de gráficos se enfoca en dos de las variables de los vuelos: los niveles de demora y las conexiones. Se muestra la ocurrencia de niveles de demoras y conexiones entre todos los vuelos para enfocar el análisis en cada uno de estos atributos 	T8																				

Representación	Codificación	Tarea																																																				
<p>Tipo de Ruta y Aeronave</p> <p>R. ● Doméstico ● Internacional</p> <table border="1"> <thead> <tr> <th>Aeronave</th> <th>Tipo de Ruta</th> <th>Vuelos</th> </tr> </thead> <tbody> <tr><td>A320</td><td>Doméstico</td><td>11 mil</td></tr> <tr><td>A321</td><td>Internacional</td><td>6 mil</td></tr> <tr><td>A319</td><td>Doméstico</td><td>2 mil</td></tr> <tr><td>B787</td><td>Internacional</td><td>4 mil</td></tr> <tr><td>B787</td><td>Doméstico</td><td>3 mil</td></tr> <tr><td>A330</td><td>Doméstico</td><td>0 mil</td></tr> <tr><td>A318</td><td>Doméstico</td><td>3 mil</td></tr> <tr><td>AT72</td><td>Doméstico</td><td>1 mil</td></tr> <tr><td>E190</td><td>Internacional</td><td>0 mil</td></tr> <tr><td>A320</td><td>Internacional</td><td>1 mil</td></tr> </tbody> </table> <p>IndMB</p> <p>Costo</p> <p>Tipos de Ruta: Doméstico (rojo), Internacional (azul)</p> <p>Aircraft Models: AT72, E190, A318, A319, A320, A330, A321, B787</p>	Aeronave	Tipo de Ruta	Vuelos	A320	Doméstico	11 mil	A321	Internacional	6 mil	A319	Doméstico	2 mil	B787	Internacional	4 mil	B787	Doméstico	3 mil	A330	Doméstico	0 mil	A318	Doméstico	3 mil	AT72	Doméstico	1 mil	E190	Internacional	0 mil	A320	Internacional	1 mil	<ul style="list-style-type: none"> Luego se visualizan con respecto a los indicadores de las irregularidades para determinar su contribución en cada uno. <p>Modismos: Horizontal Bar Chart, Scatterplot</p> <p>Marcas: Líneas, puntos (respectivamente)</p> <p>Canales (Bar chart):</p> <ul style="list-style-type: none"> Color, para identificar el tipo de ruta de los vuelos (todos los diagramas). Longitud, para identificar el número de vuelos en cada barra. Orden, de acuerdo con el número de vuelos (descendente). <p>Canales (Scatterplot):</p> <ul style="list-style-type: none"> Posición en X para el costo. Posición en Y para el indicador MB Tamaño para el número de casos Forma, para el atributo de aeronaves <p>Decisiones de diseño:</p> <ul style="list-style-type: none"> Este conjunto de gráficos se enfoca en dos de las variables de los vuelos: el tipo de ruta y el tipo de aeronave. Se muestra el volumen de vuelos realizados en cada tipo de ruta y aeronave, para enfocar el análisis en cada uno de estos atributos. Luego se visualizan con respecto a los indicadores de las irregularidades para determinar su contribución en cada uno. 	T8																			
Aeronave	Tipo de Ruta	Vuelos																																																				
A320	Doméstico	11 mil																																																				
A321	Internacional	6 mil																																																				
A319	Doméstico	2 mil																																																				
B787	Internacional	4 mil																																																				
B787	Doméstico	3 mil																																																				
A330	Doméstico	0 mil																																																				
A318	Doméstico	3 mil																																																				
AT72	Doméstico	1 mil																																																				
E190	Internacional	0 mil																																																				
A320	Internacional	1 mil																																																				
<p>Segmento, Vuelo y Ruta registrada</p> <table border="1"> <thead> <tr> <th>Segm.</th> <th>IndMB</th> <th>Costo</th> <th>Casos</th> </tr> </thead> <tbody> <tr><td>BOG-CTG</td><td>13,39</td><td>\$ 39.723</td><td>2406</td></tr> <tr><td>BOG-MDE</td><td>13,89</td><td>\$ 41.279</td><td>1964</td></tr> <tr><td>BOG-PEI</td><td>14,94</td><td>\$ 35.733</td><td>1928</td></tr> <tr><td>BOG-BAQ</td><td>12,97</td><td>\$ 30.632</td><td>1832</td></tr> <tr><td>BOG-BGA</td><td>13,39</td><td>\$ 27.813</td><td>1690</td></tr> <tr><td>BOG-CLO</td><td>14,11</td><td>\$ 7.972</td><td>1642</td></tr> <tr><td>BOG-CUC</td><td>13,88</td><td>\$ 10.509</td><td>1195</td></tr> <tr><td>BOG-SMR</td><td>12,01</td><td>\$ 15.152</td><td>1089</td></tr> <tr><td>BOG-AXM</td><td>14,07</td><td>\$ 10.198</td><td>585</td></tr> <tr><td>BOG-ADZ</td><td>13,52</td><td>\$ 7.440</td><td>536</td></tr> <tr><td>BOG-MTR</td><td>13,25</td><td>\$ 12.584</td><td>512</td></tr> <tr><td>Total</td><td>13,03</td><td>\$ 293.051</td><td>17983</td></tr> </tbody> </table>	Segm.	IndMB	Costo	Casos	BOG-CTG	13,39	\$ 39.723	2406	BOG-MDE	13,89	\$ 41.279	1964	BOG-PEI	14,94	\$ 35.733	1928	BOG-BAQ	12,97	\$ 30.632	1832	BOG-BGA	13,39	\$ 27.813	1690	BOG-CLO	14,11	\$ 7.972	1642	BOG-CUC	13,88	\$ 10.509	1195	BOG-SMR	12,01	\$ 15.152	1089	BOG-AXM	14,07	\$ 10.198	585	BOG-ADZ	13,52	\$ 7.440	536	BOG-MTR	13,25	\$ 12.584	512	Total	13,03	\$ 293.051	17983	<p>Modismo: HeatMap (matrix)</p> <p>Marcas: Áreas.</p> <p>Canales:</p> <ul style="list-style-type: none"> Un color divergente diferente para indicador MB, costos y número de casos presentados. El orden puede ser establecido por el usuario, no obstante por defecto se sitúa en el número de casos. <p>Decisiones de diseño:</p> <ul style="list-style-type: none"> Esta matriz se incluye para obtener detalles a demanda e incluso poder enfocar el análisis en algún segmento (origen-destino) en particular. 	T8
Segm.	IndMB	Costo	Casos																																																			
BOG-CTG	13,39	\$ 39.723	2406																																																			
BOG-MDE	13,89	\$ 41.279	1964																																																			
BOG-PEI	14,94	\$ 35.733	1928																																																			
BOG-BAQ	12,97	\$ 30.632	1832																																																			
BOG-BGA	13,39	\$ 27.813	1690																																																			
BOG-CLO	14,11	\$ 7.972	1642																																																			
BOG-CUC	13,88	\$ 10.509	1195																																																			
BOG-SMR	12,01	\$ 15.152	1089																																																			
BOG-AXM	14,07	\$ 10.198	585																																																			
BOG-ADZ	13,52	\$ 7.440	536																																																			
BOG-MTR	13,25	\$ 12.584	512																																																			
Total	13,03	\$ 293.051	17983																																																			



Figura 39. Vista de Condiciones Climáticas

Tabla 44. Decisiones de diseño para la vista de condiciones climáticas

Representación	Codificación	Tarea																																																																																															
<p>Irregularidades</p> <p>Responsable por # de casos</p> <table border="1"> <tr><td>BOG</td><td>64.59%</td></tr> <tr><td>MIA</td><td>3.66%</td></tr> <tr><td>MAD</td><td>3.46%</td></tr> <tr><td>SAL</td><td>3.08%</td></tr> <tr><td>LIM</td><td>2.44%</td></tr> <tr><td>MEX</td><td>1.87%</td></tr> <tr><td>0%</td><td>50%</td></tr> <tr><td>Total</td><td>10,25</td></tr> </table> <p>Casos por Segmento y Vuelo</p> <table border="1"> <thead> <tr><th>Segmento</th><th>DepDelay</th><th>ArrDelay</th><th>Casos</th><th>IndMB</th></tr> </thead> <tbody> <tr><td>BOG-MDE</td><td>9,42</td><td>8,85</td><td>1119</td><td>13,05</td></tr> <tr><td>BOG-CLO</td><td>9,61</td><td>7,84</td><td>5888</td><td>12,95</td></tr> <tr><td>BOG-CTG</td><td>6,74</td><td>4,38</td><td>3974</td><td>12,77</td></tr> <tr><td>BOG-BAQ</td><td>8,24</td><td>5,54</td><td>3373</td><td>13,51</td></tr> <tr><td>BOG-BGA</td><td>9,47</td><td>6,61</td><td>2535</td><td>14,55</td></tr> <tr><td>BOG-PEI</td><td>10,06</td><td>8,88</td><td>2512</td><td>15,65</td></tr> <tr><td>BOG-LIM</td><td>8,16</td><td>10,07</td><td>2252</td><td>15,64</td></tr> <tr><td>MAD-BOG</td><td>34,01</td><td>56,85</td><td>1987</td><td>13,36</td></tr> <tr><td>BOG-GUC</td><td>7,22</td><td>5,80</td><td>1022</td><td>11,01</td></tr> <tr><td>Total</td><td>10,25</td><td>7,92</td><td>89986</td><td>14,03</td></tr> </tbody> </table> <p>Causa asignada (RL)</p> <table border="1"> <tr><td>64-problemas.</td><td>16,27%</td></tr> <tr><td>86-daño en e...</td><td>11,21%</td></tr> <tr><td>26-equipaje r...</td><td>9,63%</td></tr> <tr><td>56-pé...</td><td>8,89%</td></tr> <tr><td>53-seg...</td><td>6,32%</td></tr> <tr><td>54-...</td><td>5,41%</td></tr> <tr><td>31-...</td><td>7,15%</td></tr> <tr><td>11-...</td><td>5,60%</td></tr> <tr><td>11-...</td><td>4,02%</td></tr> <tr><td>35-...</td><td>3,90%</td></tr> <tr><td>76-equ...</td><td>3,63%</td></tr> <tr><td>...</td><td>...</td></tr> </table>	BOG	64.59%	MIA	3.66%	MAD	3.46%	SAL	3.08%	LIM	2.44%	MEX	1.87%	0%	50%	Total	10,25	Segmento	DepDelay	ArrDelay	Casos	IndMB	BOG-MDE	9,42	8,85	1119	13,05	BOG-CLO	9,61	7,84	5888	12,95	BOG-CTG	6,74	4,38	3974	12,77	BOG-BAQ	8,24	5,54	3373	13,51	BOG-BGA	9,47	6,61	2535	14,55	BOG-PEI	10,06	8,88	2512	15,65	BOG-LIM	8,16	10,07	2252	15,64	MAD-BOG	34,01	56,85	1987	13,36	BOG-GUC	7,22	5,80	1022	11,01	Total	10,25	7,92	89986	14,03	64-problemas.	16,27%	86-daño en e...	11,21%	26-equipaje r...	9,63%	56-pé...	8,89%	53-seg...	6,32%	54-...	5,41%	31-...	7,15%	11-...	5,60%	11-...	4,02%	35-...	3,90%	76-equ...	3,63%	<p>Decisiones de diseño:</p> <ul style="list-style-type: none"> Estos tres modismos (<i>barchart</i>, <i>heatmap</i>, <i>treechart</i>) adoptan tareas de la fase descriptiva, con el fin de ubicar el punto de partida para el análisis diagnóstico. Los puntos de partida para el análisis pueden ser: el aeropuerto responsable (por número de casos), los segmentos de los vuelos (por el mayor número de demoras, casos o indicador MB) o las causas asignadas. En el caso del <i>treechart</i>, también se utiliza para comparar la causa asignada a los vuelos luego de realizar tareas de análisis sobre las demás visualizaciones. Por ejemplo, cuando se identifica un patrón en las demoras o condiciones climáticas. 	T5, T6
BOG	64.59%																																																																																																
MIA	3.66%																																																																																																
MAD	3.46%																																																																																																
SAL	3.08%																																																																																																
LIM	2.44%																																																																																																
MEX	1.87%																																																																																																
0%	50%																																																																																																
Total	10,25																																																																																																
Segmento	DepDelay	ArrDelay	Casos	IndMB																																																																																													
BOG-MDE	9,42	8,85	1119	13,05																																																																																													
BOG-CLO	9,61	7,84	5888	12,95																																																																																													
BOG-CTG	6,74	4,38	3974	12,77																																																																																													
BOG-BAQ	8,24	5,54	3373	13,51																																																																																													
BOG-BGA	9,47	6,61	2535	14,55																																																																																													
BOG-PEI	10,06	8,88	2512	15,65																																																																																													
BOG-LIM	8,16	10,07	2252	15,64																																																																																													
MAD-BOG	34,01	56,85	1987	13,36																																																																																													
BOG-GUC	7,22	5,80	1022	11,01																																																																																													
Total	10,25	7,92	89986	14,03																																																																																													
64-problemas.	16,27%																																																																																																
86-daño en e...	11,21%																																																																																																
26-equipaje r...	9,63%																																																																																																
56-pé...	8,89%																																																																																																
53-seg...	6,32%																																																																																																
54-...	5,41%																																																																																																
31-...	7,15%																																																																																																
11-...	5,60%																																																																																																
11-...	4,02%																																																																																																
35-...	3,90%																																																																																																
76-equ...	3,63%																																																																																																
...	...																																																																																																
<p>Demoras</p> <p>Demora promedio (mins) y casos por periodo</p> <p>● Demora en origen ● Demora en destino ● Casos</p> <p>Demora prom.</p> <p>Casos</p> <p>enero febrero marzo abril mayo junio julio agosto septiembre octubre</p> <p>20 mil</p> <p>10 mil</p>	<p>Modismo: <i>Line Chart</i></p> <p>Marcas: Puntos con marcas de conexión (líneas).</p> <p>Canales:</p> <ul style="list-style-type: none"> Color, para identificar cada atributo (demora en origen, destino y número de casos). Orden, ascendente por variable temporal (mes, día). Se utilizan etiquetas adicionales al ubicarse sobre cada categoría para ver el promedio de la demora y el número de casos. <p>Decisiones de diseño:</p> <ul style="list-style-type: none"> Se usan los atributos de Demoras en Vuelos y Casos, para apoyar la tarea de identificar correlación entre estos dos atributos en un momento dado. No se incluye el indicador MB en este mismo diagrama, ya que su escala es muy diferente a la de estos dos atributos y en la representación no es perceptible (se ubica junto con el eje horizontal de tiempo). 	T8																																																																																															

Representación	Codificación	Tarea
<p>Demora promedio en el origen por hora del día</p> <p>Nivel Demora (mins) ● < 15 ● > 45 ● 15-45</p> <p>Demora Promedio por hora del día</p> <p>Demora promedio en el destino por hora del día</p> <p>Nivel Demora (mins) ● < 15 ● > 45 ● 15-45</p> <p>Demora Promedio por hora del día</p>	<p>Modismo: Vertical Stacked Barchart Marcas: Líneas. Canales:</p> <ul style="list-style-type: none"> Longitud, para indicar el valor de la demora promedio en el vuelo. Color, para identificar el nivel de demora. Posición en X, para representar cada hora del día. Orden, de acuerdo con la variable temporal (horas) <p>Decisiones de diseño:</p> <ul style="list-style-type: none"> Se utilizan barras apiladas y no barras paralelas para que el diagrama pueda escalar mejor. Se utiliza este diagrama con base en las horas del día, ya que las condiciones climáticas también están expresadas en un momento del día, y el análisis de correlación requiere que se encuentren en la misma escala de medición. Este diagrama no incorpora los indicadores de las irregularidades. Se utiliza como punto de ubicación del momento específico del día en que se quiere profundizar el análisis, ya sea de manera general (iniciando desde aquí) o de manera detallada (usando las demás visualizaciones como filtro hasta llegar aquí). 	T8
<p>Visibilidad promedio en estación origen por hora del día</p> <p>Visibilidad promedio en estación destino por hora del día</p> <p>Visibilidad promedio por hora del día</p>	<p>Modismo: Area Chart Marcas: Áreas. Canales:</p> <ul style="list-style-type: none"> Color, para identificar el atributo (visibilidad). Orden, ascendente por variable temporal (hora del día). <p>Decisiones de diseño:</p> <ul style="list-style-type: none"> Este diagrama inicia con una vista agregada (promedio) de visibilidad por hora del día. Sin embargo, su objetivo es utilizar las demás visualizaciones para ubicar el foco de análisis y llegar a un día específico. Se incluye un diagrama por cada extremo de la ruta (origen y destino), ya que las demoras generadas en los vuelos pueden atribuirse al origen, al destino o a ambos, y es importante poder analizarlos por separado. 	T9

Representación	Codificación	Tarea																																				
<p>Nubosidad en origen</p> <table border="1"> <thead> <tr> <th>Nubosidad</th> <th>Porcentaje</th> </tr> </thead> <tbody> <tr> <td>FEW</td> <td>42.59%</td> </tr> <tr> <td>SCT</td> <td>32.32%</td> </tr> <tr> <td>BKN</td> <td>12.02%</td> </tr> <tr> <td>CLR</td> <td>10.81%</td> </tr> </tbody> </table> <p>Reglas de Vuelo (origen)</p> <table border="1"> <thead> <tr> <th>Regla de Vuelo</th> <th>Porcentaje</th> </tr> </thead> <tbody> <tr> <td>VFR</td> <td>72.55%</td> </tr> <tr> <td>MVFR</td> <td>22.9%</td> </tr> <tr> <td>IFR</td> <td>2.5%</td> </tr> </tbody> </table> <p>Nubosidad en destino</p> <table border="1"> <thead> <tr> <th>Nubosidad</th> <th>Porcentaje</th> </tr> </thead> <tbody> <tr> <td>FEW</td> <td>42.59%</td> </tr> <tr> <td>SCT</td> <td>32.32%</td> </tr> <tr> <td>BKN</td> <td>12.02%</td> </tr> <tr> <td>CLR</td> <td>10.81%</td> </tr> </tbody> </table> <p>Reglas de Vuelo (destino)</p> <table border="1"> <thead> <tr> <th>Regla de Vuelo</th> <th>Porcentaje</th> </tr> </thead> <tbody> <tr> <td>VFR</td> <td>73.63%</td> </tr> <tr> <td>MVFR</td> <td>18.7%</td> </tr> <tr> <td>IFR</td> <td>7.2%</td> </tr> </tbody> </table>	Nubosidad	Porcentaje	FEW	42.59%	SCT	32.32%	BKN	12.02%	CLR	10.81%	Regla de Vuelo	Porcentaje	VFR	72.55%	MVFR	22.9%	IFR	2.5%	Nubosidad	Porcentaje	FEW	42.59%	SCT	32.32%	BKN	12.02%	CLR	10.81%	Regla de Vuelo	Porcentaje	VFR	73.63%	MVFR	18.7%	IFR	7.2%	<p>Modismos: Horizontal Bar Chart, Donut chart</p> <p>Marcas: líneas, áreas (respectivamente).</p> <p>Canales (barchart):</p> <ul style="list-style-type: none"> Color divergente, para identificar el tipo de nubosidad en el sitio (aeropuerto de origen o destino). Más oscuro indica mayor presencia de nubes (y por tanto menor visibilidad). Longitud, para identificar el número de lecturas de clima que presentó cada tipo de nubosidad (en proporción). Orden, de acuerdo con el número de lecturas de clima en cada tipo (descendente). <p>Canales (donut chart):</p> <ul style="list-style-type: none"> Ángulo, con distribución radial. Color, para el tipo de regla de vuelo. <p>Decisiones de diseño:</p> <ul style="list-style-type: none"> Este diagrama inicia con una vista agregada (conteo) de lecturas por cada tipo de nubes o reglas de vuelo. Sin embargo, su objetivo es utilizar las demás visualizaciones para ubicar el foco de análisis y llegar a un día específico. Se incluye un diagrama por cada extremo de la ruta (origen y destino), para poder analizarlas en relación con las demoras de los vuelos en origen y destino. También es posible utilizar estas visualizaciones como punto de partida para analizar todos los vuelos cuyas condiciones en origen o destino presentaron condiciones climáticas adversas (demasiadas nubes, poca visibilidad, reglas de vuelo basadas en instrumentos) y ver los niveles de demora y las irregularidades presentadas (número de casos, indicador MB) en esos vuelos. 	T8
Nubosidad	Porcentaje																																					
FEW	42.59%																																					
SCT	32.32%																																					
BKN	12.02%																																					
CLR	10.81%																																					
Regla de Vuelo	Porcentaje																																					
VFR	72.55%																																					
MVFR	22.9%																																					
IFR	2.5%																																					
Nubosidad	Porcentaje																																					
FEW	42.59%																																					
SCT	32.32%																																					
BKN	12.02%																																					
CLR	10.81%																																					
Regla de Vuelo	Porcentaje																																					
VFR	73.63%																																					
MVFR	18.7%																																					
IFR	7.2%																																					

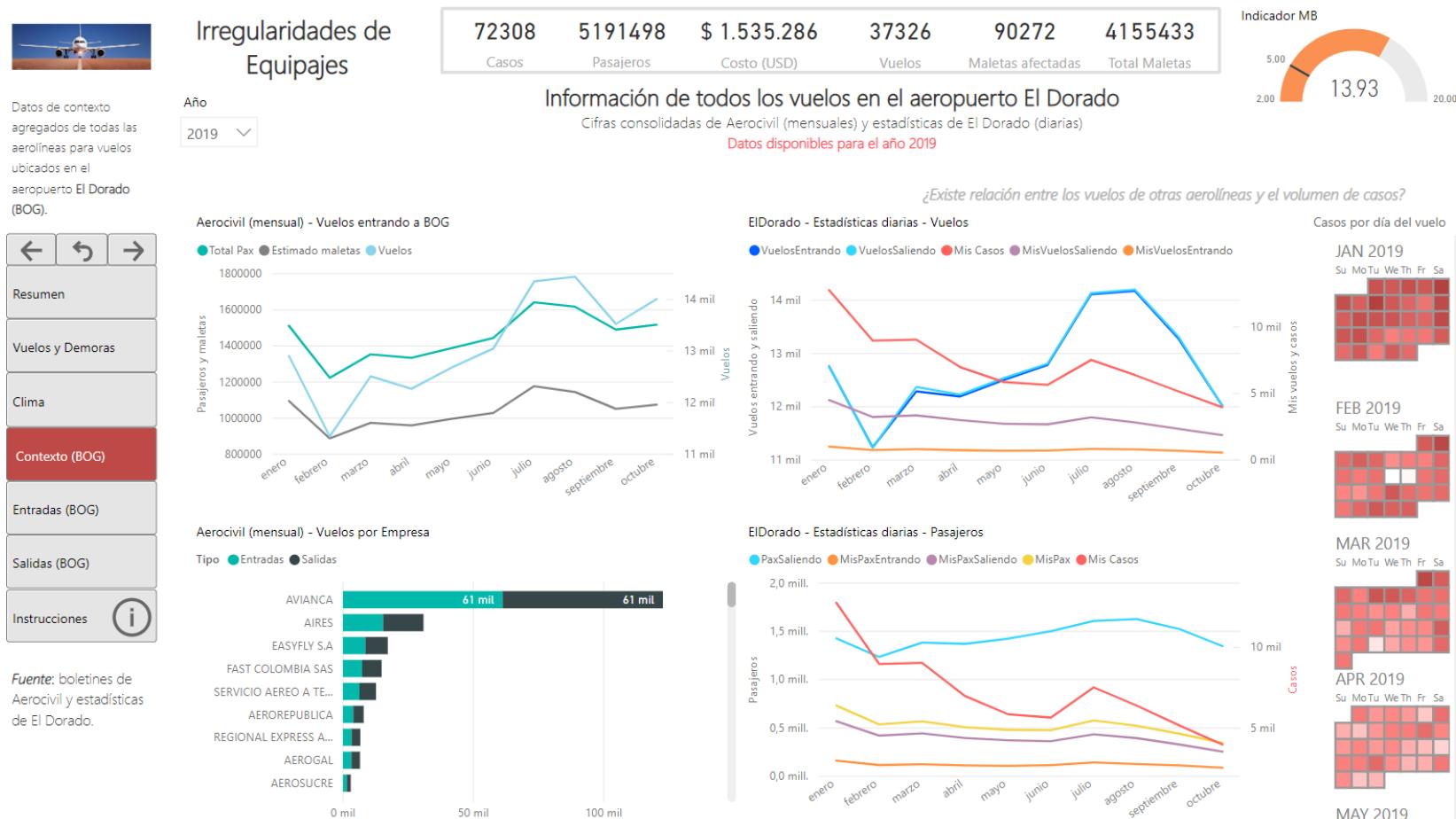
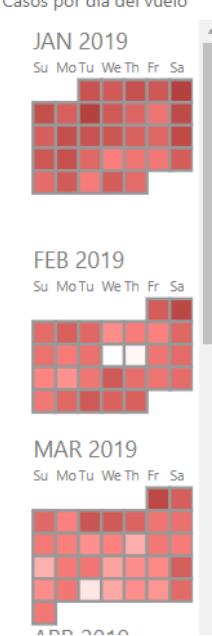


Figura 40. Vista de Contexto

Tabla 45. Decisiones de diseño para la vista de contexto

Representación	Codificación	Tarea
<p>Aerocivil (mensual) - Vuelos entrando a BOG</p> <p>Total Pax ● Estimado maletas ● Vuelos</p> <p>Pasajeros y maletas</p> <p>Vuelos</p> <p>enero febrero marzo abril mayo junio julio agosto septiembre octubre</p> <p>Aerocivil (mensual) - Vuelos por Empresa</p> <p>Tipo ● Entradas ● Salidas</p> <p>AVIANCA 61 mil 61 mil</p> <p>AIRES</p> <p>EASYFLY S.A.</p> <p>FAST COLOMBIA SAS</p> <p>SERVICIO AEREO A TEL...</p> <p>AEROREPUBLICA</p> <p>REGIONAL EXPRESS A...</p> <p>AEROGAL</p> <p>AEROSUCRE</p> <p>0 mil 50 mil 100 mil</p>	<p>Modismo: Line Chart; Horizontal Stacked Barchart</p> <p>Marcas: Puntos con marcas de conexión (líneas); líneas (respectivamente).</p> <p>Canales (line chart):</p> <ul style="list-style-type: none"> Color, para identificar cada atributo (pasajeros, maletas, vuelos). Orden, ascendente por variable temporal (mes, día). Se utilizan etiquetas adicionales al ubicarse sobre cada categoría para ver el valor exacto de cada atributo en un momento dato. <p>Canales (bar chart):</p> <ul style="list-style-type: none"> Longitud, para indicar el volumen de vuelos realizado por cada aerolínea. Color, para identificar si el vuelo ingresa o sale del aeropuerto El Dorado. Posición en Y y orden descendente según el número de vuelos realizados. <p>Decisiones de diseño:</p> <ul style="list-style-type: none"> Se utiliza esta información de la fuente <i>Aerocivil</i> como contexto para poder comparar contra los vuelos realizados por la aerolínea en el caso de estudio. Esta información tiene distinto nivel de agregación (mensual), por tanto no se coordina con las demás visualizaciones. 	T10
<p>El Dorado - Estadísticas diarias - Vuelos</p> <p>VuelosEntrando ● VuelosSaliendo ● Mis Casos ● MisVuelosSaliente ● MisVuelosEntrando</p> <p>Mil vuelos y casos</p> <p>enero febrero marzo abril mayo junio julio agosto septiembre octubre</p> <p>El Dorado - Estadísticas diarias - Pasajeros</p> <p>PaxSaliente ● PaxEntrando ● MisPaxSaliente ● MisPaxEntrando ● MisCasos</p> <p>Pasajeros</p> <p>Casos</p> <p>enero febrero marzo abril mayo junio julio agosto septiembre octubre</p>	<p>Modismo: Line Chart (un diagrama para vuelos y otro para pasajeros)</p> <p>Marcas: Puntos con marcas de conexión (líneas).</p> <p>Canales:</p> <ul style="list-style-type: none"> Color, para identificar cada atributo (vuelos, pasajeros, casos). Orden, ascendente por variable temporal (mes, día). Se utilizan etiquetas adicionales al ubicarse sobre cada categoría para ver el volumen de vuelos, pasajeros y casos. <p>Decisiones de diseño:</p> <ul style="list-style-type: none"> Se emplean dos diagramas, uno para vuelos y otro para pasajeros, ya que manejan dimensiones diferentes y por tanto requieren escalas diferentes que no pueden ser representadas en un mismo diagrama sin afectación en la percepción. Se diferencian los vuelos de otras aerolíneas de la aerolínea del caso de estudio, también sucede lo mismo con los pasajeros. No se incluye el indicador MB en este mismo diagrama, ya que su escala es muy diferente a la de estos dos atributos y en la representación no es perceptible (se ubica junto con el eje horizontal de tiempo). 	T10

Representación	Codificación	Tarea
<p>Casos por día del vuelo</p>  <p>JAN 2019 Su Mo Tu We Th Fr Sa</p> <p>FEB 2019 Su Mo Tu We Th Fr Sa</p> <p>MAR 2019 Su Mo Tu We Th Fr Sa</p> <p>APR 2019</p>	<p>Modismo: HeatMap (<i>calendar</i>)</p> <p>Marcas: Áreas.</p> <p>Canales:</p> <ul style="list-style-type: none"> ▪ Un color divergente para indicador MB, representando días con mayor número de casos en tono más oscuro. ▪ El orden está establecido por la disposición de los datos en el calendario (temporal). <p>Decisiones de diseño:</p> <ul style="list-style-type: none"> ▪ Esta matriz se incluye para situar el foco del análisis en un mes o día particular. ▪ No se encuentra coordinada con las demás visualizaciones por los diferentes niveles de agregación de los datos. <p>Las 3 visualizaciones utilizadas en esta vista están enfocadas en el nivel <i>Overview</i> del <i>visualization mantra</i>, considerando la agregación de los datos, pero permitiendo una vista general del contexto en términos de vuelos y pasajeros contra las otras aerolíneas.</p>	T10

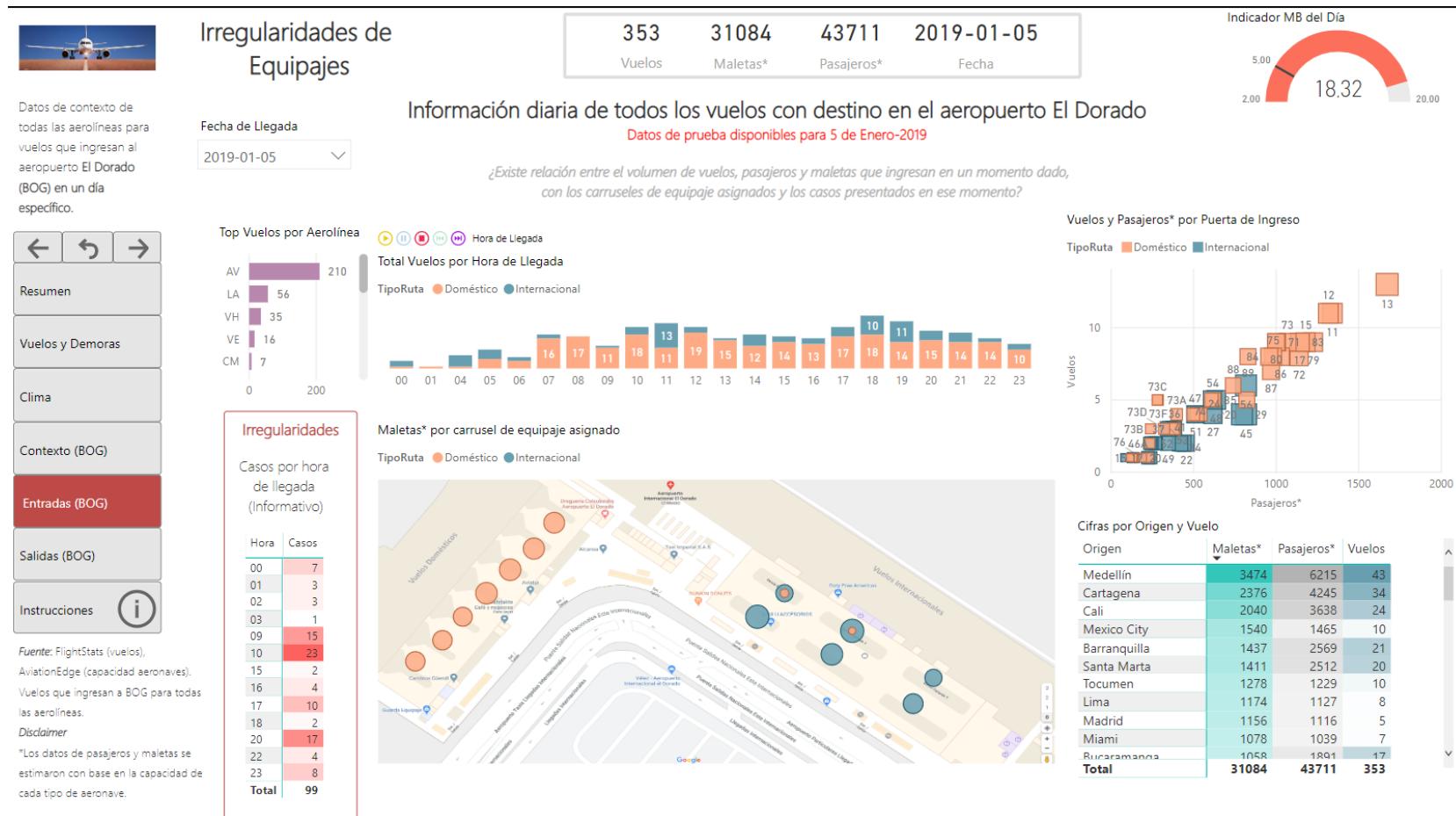
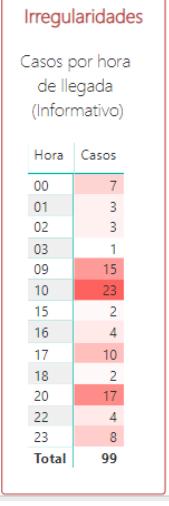
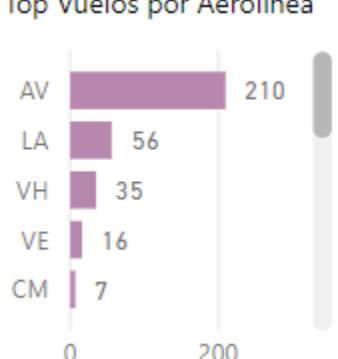


Figura 41. Vista de Contexto - Vuelos ingresando (entradas) a El Dorado

Tabla 46. Decisiones de diseño para la vista de entradas a El Dorado

Representación	Codificación	Tarea																														
 <p>Irregularidades Casos por hora de llegada (Informativo)</p> <table border="1"> <thead> <tr> <th>Hora</th> <th>Casos</th> </tr> </thead> <tbody> <tr><td>00</td><td>7</td></tr> <tr><td>01</td><td>3</td></tr> <tr><td>02</td><td>3</td></tr> <tr><td>03</td><td>1</td></tr> <tr><td>09</td><td>15</td></tr> <tr><td>10</td><td>23</td></tr> <tr><td>15</td><td>2</td></tr> <tr><td>16</td><td>4</td></tr> <tr><td>17</td><td>10</td></tr> <tr><td>18</td><td>2</td></tr> <tr><td>20</td><td>17</td></tr> <tr><td>22</td><td>4</td></tr> <tr><td>23</td><td>8</td></tr> <tr><td>Total</td><td>99</td></tr> </tbody> </table>	Hora	Casos	00	7	01	3	02	3	03	1	09	15	10	23	15	2	16	4	17	10	18	2	20	17	22	4	23	8	Total	99	Modismo: HeatMap (matrix) Marcas: Áreas. Canales: <ul style="list-style-type: none"> Un color divergente para los casos presentados. El orden puede ser establecido por el usuario, no obstante por defecto se sitúa en la hora del día. Decisiones de diseño: <ul style="list-style-type: none"> Esta matriz se incluye para enfocar el análisis en un momento específico del día, a partir de los casos presentados en ese momento. No presenta coordinación con las demás vistas. 	T11
Hora	Casos																															
00	7																															
01	3																															
02	3																															
03	1																															
09	15																															
10	23																															
15	2																															
16	4																															
17	10																															
18	2																															
20	17																															
22	4																															
23	8																															
Total	99																															
 <p>Top Vuelos por Aerolínea</p> <table> <thead> <tr> <th>Aerolínea</th> <th>Vuelos</th> </tr> </thead> <tbody> <tr><td>AV</td><td>210</td></tr> <tr><td>LA</td><td>56</td></tr> <tr><td>VH</td><td>35</td></tr> <tr><td>VE</td><td>16</td></tr> <tr><td>CM</td><td>7</td></tr> </tbody> </table>	Aerolínea	Vuelos	AV	210	LA	56	VH	35	VE	16	CM	7	Modismo: Horizontal Bar Chart Marcas: líneas Canales: <ul style="list-style-type: none"> Longitud, para identificar el número de vuelos en cada barra. Posición en Y y Orden descendente, de acuerdo con el número de vuelos realizados por cada aerolínea. Decisiones de diseño: <ul style="list-style-type: none"> Se incluye este diagrama para obtener un entendimiento de la predominancia de los vuelos de una aerolínea en particular. Se utiliza como filtro si se desea enfocar el análisis en una compañía específica. También como destino del análisis si, luego de haberse enfocado previamente en un momento del día, se desea conocer si la congestión en alguna zona del aeropuerto se debió a vuelos de una compañía en particular. 	T11																		
Aerolínea	Vuelos																															
AV	210																															
LA	56																															
VH	35																															
VE	16																															
CM	7																															

<p>Total Vuelos por Hora de Llegada</p> <p>TipoRuta Doméstico Internacional</p> <table border="1"> <thead> <tr> <th>Hora</th> <th>Doméstico</th> <th>Internacional</th> <th>Total</th> </tr> </thead> <tbody> <tr><td>00</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>01</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>04</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>05</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>06</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>07</td><td>16</td><td>0</td><td>16</td></tr> <tr><td>08</td><td>17</td><td>0</td><td>17</td></tr> <tr><td>09</td><td>11</td><td>0</td><td>11</td></tr> <tr><td>10</td><td>18</td><td>13</td><td>31</td></tr> <tr><td>11</td><td>11</td><td>13</td><td>24</td></tr> <tr><td>12</td><td>19</td><td>12</td><td>31</td></tr> <tr><td>13</td><td>15</td><td>14</td><td>29</td></tr> <tr><td>14</td><td>12</td><td>13</td><td>25</td></tr> <tr><td>15</td><td>14</td><td>13</td><td>27</td></tr> <tr><td>16</td><td>13</td><td>10</td><td>23</td></tr> <tr><td>17</td><td>17</td><td>10</td><td>27</td></tr> <tr><td>18</td><td>18</td><td>1</td><td>19</td></tr> </tbody> </table>	Hora	Doméstico	Internacional	Total	00	0	0	0	01	0	0	0	04	0	0	0	05	0	0	0	06	0	0	0	07	16	0	16	08	17	0	17	09	11	0	11	10	18	13	31	11	11	13	24	12	19	12	31	13	15	14	29	14	12	13	25	15	14	13	27	16	13	10	23	17	17	10	27	18	18	1	19	<p>Modismo: Vertical Stacked Barchart Marcas: Líneas. Canales:</p> <ul style="list-style-type: none"> Longitud, para indicar el número de vuelos. Color, para identificar el tipo de ruta (doméstica, internacional). Posición en X, para representar cada hora del día. Orden, de acuerdo con la variable temporal (horas) <p>Decisiones de diseño:</p> <ul style="list-style-type: none"> Se utilizan barras apiladas y no barras paralelas para que el diagrama pueda escalar mejor. Se utiliza este diagrama con base en las horas del día, ya que la vista completa está diseñada para el análisis de un día específico. Este diagrama no incorpora los indicadores de las irregularidades. Se utiliza como punto de ubicación del momento específico del día en que se quiere profundizar el análisis (a través de la localización del momento específico de un pico en el número de vuelos que ingresaron). 	T11																											
Hora	Doméstico	Internacional	Total																																																																																																		
00	0	0	0																																																																																																		
01	0	0	0																																																																																																		
04	0	0	0																																																																																																		
05	0	0	0																																																																																																		
06	0	0	0																																																																																																		
07	16	0	16																																																																																																		
08	17	0	17																																																																																																		
09	11	0	11																																																																																																		
10	18	13	31																																																																																																		
11	11	13	24																																																																																																		
12	19	12	31																																																																																																		
13	15	14	29																																																																																																		
14	12	13	25																																																																																																		
15	14	13	27																																																																																																		
16	13	10	23																																																																																																		
17	17	10	27																																																																																																		
18	18	1	19																																																																																																		
<p>Vuelos y Pasajeros* por Puerta de Ingreso</p> <p>TipoRuta Doméstico Internacional</p> <table border="1"> <thead> <tr> <th>Puerta</th> <th>Pasajeros*</th> <th>Vuelos</th> </tr> </thead> <tbody> <tr><td>73C</td><td>300</td><td>5</td></tr> <tr><td>73A</td><td>400</td><td>4</td></tr> <tr><td>73D</td><td>200</td><td>3</td></tr> <tr><td>73F</td><td>350</td><td>6</td></tr> <tr><td>73B</td><td>250</td><td>4</td></tr> <tr><td>76</td><td>150</td><td>2</td></tr> <tr><td>46A</td><td>200</td><td>3</td></tr> <tr><td>18</td><td>100</td><td>2</td></tr> <tr><td>73C</td><td>450</td><td>54</td></tr> <tr><td>73A</td><td>500</td><td>54</td></tr> <tr><td>73D</td><td>400</td><td>47</td></tr> <tr><td>73F</td><td>550</td><td>74</td></tr> <tr><td>73B</td><td>350</td><td>48</td></tr> <tr><td>76</td><td>300</td><td>51</td></tr> <tr><td>46A</td><td>350</td><td>27</td></tr> <tr><td>18</td><td>200</td><td>45</td></tr> <tr><td>73C</td><td>600</td><td>88</td></tr> <tr><td>73A</td><td>700</td><td>89</td></tr> <tr><td>73D</td><td>550</td><td>84</td></tr> <tr><td>73F</td><td>750</td><td>80</td></tr> <tr><td>73B</td><td>650</td><td>75</td></tr> <tr><td>76</td><td>500</td><td>71</td></tr> <tr><td>46A</td><td>600</td><td>79</td></tr> <tr><td>18</td><td>400</td><td>86</td></tr> <tr><td>73C</td><td>800</td><td>12</td></tr> <tr><td>73A</td><td>900</td><td>11</td></tr> <tr><td>73D</td><td>750</td><td>72</td></tr> <tr><td>73F</td><td>950</td><td>77</td></tr> <tr><td>73B</td><td>850</td><td>79</td></tr> <tr><td>76</td><td>700</td><td>87</td></tr> <tr><td>46A</td><td>800</td><td>80</td></tr> <tr><td>18</td><td>600</td><td>83</td></tr> </tbody> </table>	Puerta	Pasajeros*	Vuelos	73C	300	5	73A	400	4	73D	200	3	73F	350	6	73B	250	4	76	150	2	46A	200	3	18	100	2	73C	450	54	73A	500	54	73D	400	47	73F	550	74	73B	350	48	76	300	51	46A	350	27	18	200	45	73C	600	88	73A	700	89	73D	550	84	73F	750	80	73B	650	75	76	500	71	46A	600	79	18	400	86	73C	800	12	73A	900	11	73D	750	72	73F	950	77	73B	850	79	76	700	87	46A	800	80	18	600	83	<p>Modismo: Scatterplot Marcas: Puntos. Canales (Scatterplot):</p> <ul style="list-style-type: none"> Color, para identificar el tipo de ruta de los vuelos Posición en X para el número de pasajeros Posición en Y para el número de vuelos que ingresan Tamaño para el número de maletas Forma, para el atributo de puertas de ingreso <p>Decisiones de diseño:</p> <ul style="list-style-type: none"> Se agregan las variables de vuelos y pasajeros ingresando, a través de esta gráfica, para contrastarla con las irregularidades presentadas en un momento específico del día. Por ejemplo, para determinar si se presentó congestión en las puertas en un momento dado y esto pudo ocasionar demoras en el ingreso. 	T11
Puerta	Pasajeros*	Vuelos																																																																																																			
73C	300	5																																																																																																			
73A	400	4																																																																																																			
73D	200	3																																																																																																			
73F	350	6																																																																																																			
73B	250	4																																																																																																			
76	150	2																																																																																																			
46A	200	3																																																																																																			
18	100	2																																																																																																			
73C	450	54																																																																																																			
73A	500	54																																																																																																			
73D	400	47																																																																																																			
73F	550	74																																																																																																			
73B	350	48																																																																																																			
76	300	51																																																																																																			
46A	350	27																																																																																																			
18	200	45																																																																																																			
73C	600	88																																																																																																			
73A	700	89																																																																																																			
73D	550	84																																																																																																			
73F	750	80																																																																																																			
73B	650	75																																																																																																			
76	500	71																																																																																																			
46A	600	79																																																																																																			
18	400	86																																																																																																			
73C	800	12																																																																																																			
73A	900	11																																																																																																			
73D	750	72																																																																																																			
73F	950	77																																																																																																			
73B	850	79																																																																																																			
76	700	87																																																																																																			
46A	800	80																																																																																																			
18	600	83																																																																																																			

 <p>Maletas* por carrusel de equipaje asignado</p> <p>TipoRuta ● Doméstico ● Internacional</p>	<p>Modismo: Scatterplot (map)</p> <p>Marcas: Puntos.</p> <p>Canales:</p> <ul style="list-style-type: none"> Coordenadas en X y Y para la ubicación espacial del carrusel de equipaje según el mapa. Color, para identificar el tipo de ruta de los vuelos (doméstica, internacional). Tamaño para el número estimado de maletas que ingresaron <p>Decisiones de diseño:</p> <ul style="list-style-type: none"> Se utiliza el scatterplot sobre el mapa real (Google, 2020b) de la zona de cáruseles de equipajes para utilizar las características preatentivas del usuario en relación con su conocimiento de la ubicación de cada punto, considerando las implicaciones que puede tener la congestión de un carrusel en un momento dado respecto del entorno de esa zona del aeropuerto. Se incluye el tipo de ruta considerando la distribución de los cáruseles de equipaje en el mapa: existe una zona para vuelos internacionales y otra para vuelos domésticos. Esto permite identificar casos de asignación de cáruseles a vuelos en una zona incorrecta. El tamaño de los puntos con base en las maletas permite conocer si en un momento del día hubo asignación en exceso a un carrusel en específico. La distribución en el mapa también permite conocer si hubo cáruseles sin asignación, o potencialmente fuera de servicio. 	T11																																																				
<p>Cifras por Origen y Vuelo</p> <table border="1"> <thead> <tr> <th>Origen</th> <th>Maletas*</th> <th>Pasajeros*</th> <th>Vuelos</th> </tr> </thead> <tbody> <tr> <td>Medellín</td> <td>3474</td> <td>6215</td> <td>43</td> </tr> <tr> <td>Cartagena</td> <td>2376</td> <td>4245</td> <td>34</td> </tr> <tr> <td>Cali</td> <td>2040</td> <td>3638</td> <td>24</td> </tr> <tr> <td>Mexico City</td> <td>1540</td> <td>1465</td> <td>10</td> </tr> <tr> <td>Barranquilla</td> <td>1437</td> <td>2569</td> <td>21</td> </tr> <tr> <td>Santa Marta</td> <td>1411</td> <td>2512</td> <td>20</td> </tr> <tr> <td>Tocumen</td> <td>1278</td> <td>1229</td> <td>10</td> </tr> <tr> <td>Lima</td> <td>1174</td> <td>1127</td> <td>8</td> </tr> <tr> <td>Madrid</td> <td>1156</td> <td>1116</td> <td>5</td> </tr> <tr> <td>Miami</td> <td>1078</td> <td>1039</td> <td>7</td> </tr> <tr> <td>Ricardamente</td> <td>1058</td> <td>1891</td> <td>17</td> </tr> <tr> <td>Total</td> <td>31084</td> <td>43711</td> <td>353</td> </tr> </tbody> </table>	Origen	Maletas*	Pasajeros*	Vuelos	Medellín	3474	6215	43	Cartagena	2376	4245	34	Cali	2040	3638	24	Mexico City	1540	1465	10	Barranquilla	1437	2569	21	Santa Marta	1411	2512	20	Tocumen	1278	1229	10	Lima	1174	1127	8	Madrid	1156	1116	5	Miami	1078	1039	7	Ricardamente	1058	1891	17	Total	31084	43711	353	<p>Modismo: HeatMap (matrix)</p> <p>Marcas: Áreas.</p> <p>Canales:</p> <ul style="list-style-type: none"> Un color divergente para cada atributo (maletas, pasajeros, vuelos). Tono más oscuro para los valores más altos en cada caso. El orden puede ser establecido por el usuario, no obstante por defecto se sitúa en las maletas, que es el principal foco de análisis en esta vista. <p>Decisiones de diseño:</p> <ul style="list-style-type: none"> Esta matriz se incluye para apoyar el análisis de datos a demanda, agregando como dato de análisis el origen desde el cual provienen los vuelos, así como el vuelo específico. Está coordinada con las demás vistas, de tal manera que puede ser usada también como punto de partida del análisis o para complementar filtros desde otros diagramas. 	T11
Origen	Maletas*	Pasajeros*	Vuelos																																																			
Medellín	3474	6215	43																																																			
Cartagena	2376	4245	34																																																			
Cali	2040	3638	24																																																			
Mexico City	1540	1465	10																																																			
Barranquilla	1437	2569	21																																																			
Santa Marta	1411	2512	20																																																			
Tocumen	1278	1229	10																																																			
Lima	1174	1127	8																																																			
Madrid	1156	1116	5																																																			
Miami	1078	1039	7																																																			
Ricardamente	1058	1891	17																																																			
Total	31084	43711	353																																																			

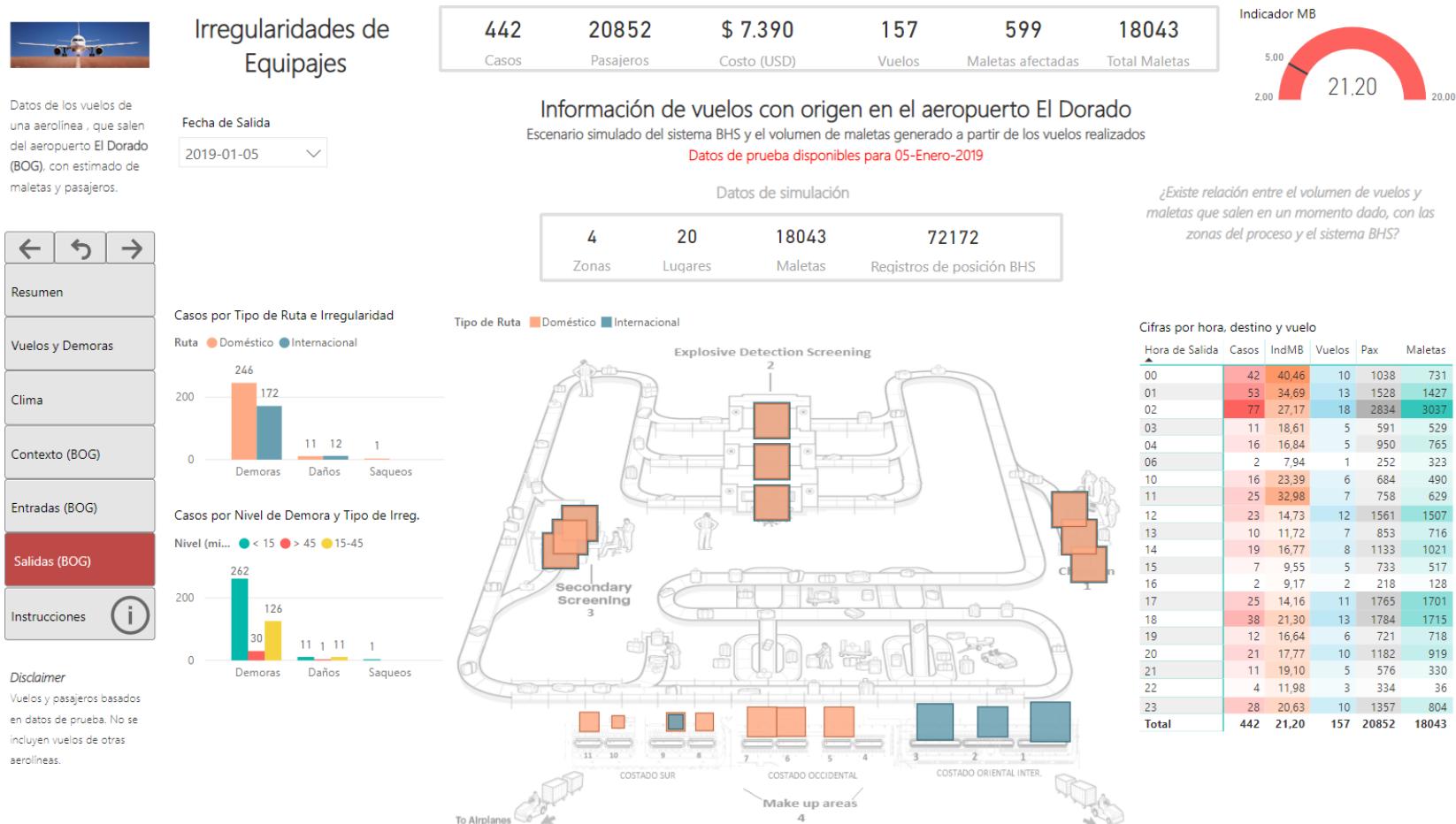
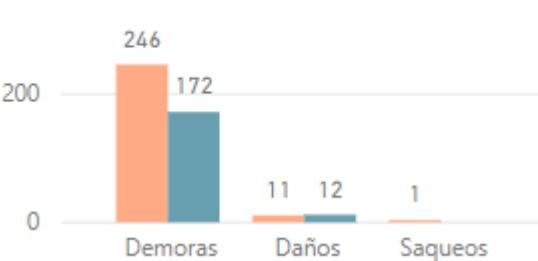
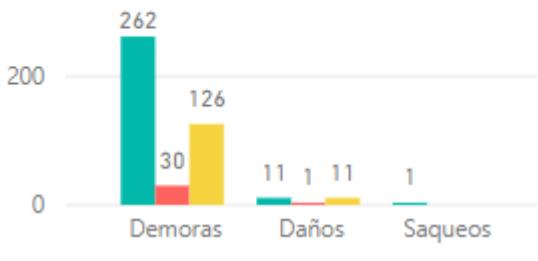
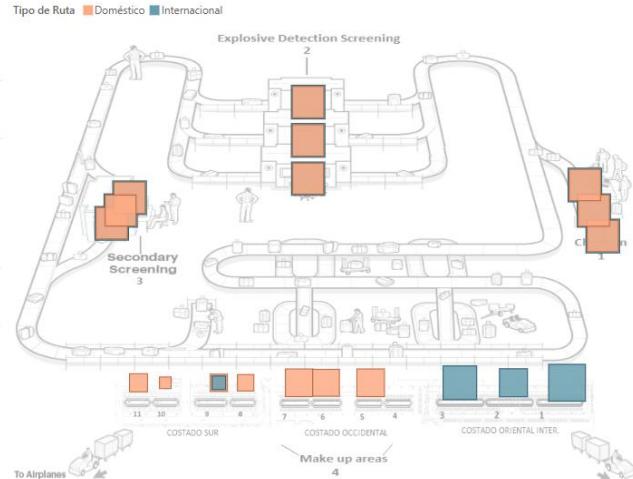


Figura 42. Vista de Contexto - Vuelos saliendo de El Dorado

Tabla 47. Decisiones de diseño para la vista de salidas

Representación	Codificación	Tarea																
<p>Casos por Tipo de Ruta e Irregularidad</p> <p>Ruta ● Doméstico ● Internacional</p>  <table border="1"> <thead> <tr> <th>Categoría</th> <th>Doméstico</th> <th>Internacional</th> </tr> </thead> <tbody> <tr> <td>Demoras</td> <td>246</td> <td>172</td> </tr> <tr> <td>Daños</td> <td>11</td> <td>12</td> </tr> <tr> <td>Saqueos</td> <td>1</td> <td>-</td> </tr> </tbody> </table>	Categoría	Doméstico	Internacional	Demoras	246	172	Daños	11	12	Saqueos	1	-	<p>Modismo: Vertical Bar Chart</p> <p>Marcas: líneas</p> <p>Canales:</p> <ul style="list-style-type: none"> Longitud, para representar el número de casos presentados. Posición en X y Orden descendente, de acuerdo con el número de casos presentados por cada tipo de irregularidad. Color, para el tipo de ruta (doméstica, internacional) <p>Decisiones de diseño:</p> <ul style="list-style-type: none"> Se incluye este diagrama para obtener un entendimiento de la predominancia de los casos presentados por tipo de irregularidad y ruta del vuelo. Se utiliza como filtro si se desea enfocar el análisis en uno de estos dos atributos o ambos (las demoras en vuelos internacionales, por ejemplo). También se utiliza como destino del análisis si, luego de haberse enfocado previamente en un momento del día, se desea conocer si la congestión en alguna zona del aeropuerto se debió a un tipo de ruta en particular, o si en ese momento se generó un pico en algún tipo de irregularidad en particular. En este caso como el número de tipos de rutas e irregularidades es bajo, se usan las barras alineadas y no apiladas, para poder realizar mejor la comparación entre cada barra y atributo. 	T12				
Categoría	Doméstico	Internacional																
Demoras	246	172																
Daños	11	12																
Saqueos	1	-																
<p>Casos por Nivel de Demora y Tipo de Irreg.</p> <p>Nivel (mi... ● < 15 ● > 45 ● 15-45</p>  <table border="1"> <thead> <tr> <th>Categoría</th> <th>< 15</th> <th>> 45</th> <th>15-45</th> </tr> </thead> <tbody> <tr> <td>Demoras</td> <td>262</td> <td>30</td> <td>126</td> </tr> <tr> <td>Daños</td> <td>11</td> <td>1</td> <td>11</td> </tr> <tr> <td>Saqueos</td> <td>1</td> <td>-</td> <td>-</td> </tr> </tbody> </table>	Categoría	< 15	> 45	15-45	Demoras	262	30	126	Daños	11	1	11	Saqueos	1	-	-	<p>Modismo: Vertical Bar Chart</p> <p>Marcas: líneas</p> <p>Canales:</p> <ul style="list-style-type: none"> Longitud, para representar el número de casos presentados. Posición en X y Orden descendente, de acuerdo con el número de casos presentados por cada nivel de demora del vuelo. Color, para el nivel de demora en la salida del vuelo (< 15 mins, entre 15 y 45 mins, > 45 mins). <p>Decisiones de diseño:</p> <ul style="list-style-type: none"> Es importante destacar que en esta vista el nivel de demora que interesa es el de salida de los vuelos y no el de llegada al destino. Uno de los propósitos del análisis es correlacionar si los vuelos que llevan demora en la salida están asociados con casos presentados en un momento específico del día o con algún tipo de irregularidad en particular. Se utiliza como filtro si se desea enfocar el análisis en uno de estos dos atributos o ambos (los daños en vuelos con nivel de demora < a 15 mins, por ejemplo). 	
Categoría	< 15	> 45	15-45															
Demoras	262	30	126															
Daños	11	1	11															
Saqueos	1	-	-															

Representación	Codificación	Tarea
	<ul style="list-style-type: none"> También se utiliza como destino del análisis si, luego de haberse enfocado previamente en un momento del día, se desea conocer si la congestión en alguna zona del aeropuerto se debió a un nivel de demora en particular, o si en ese momento se generó un pico en algún tipo de irregularidad en particular. En este caso como el número de niveles de demora y de irregularidades es bajo, se usan las barras alineadas y no apiladas, para poder realizar mejor la comparación entre cada barra y atributo. 	
 <p>Tipo de Ruta ■ Doméstico ■ Internacional</p> <p>Explosive Detection Screening 2</p> <p>Secondary Screening 3</p> <p>To Airplanes 4</p> <p>COSTADO SUR COSTADO OCCIDENTAL COSTADO ORIENTAL INTER.</p> <p>Make up areas</p>	<p>Modismo: Scatterplot (map)</p> <p>Marcas: Puntos.</p> <p>Canales:</p> <ul style="list-style-type: none"> Coordenadas en X y Y para la localización espacial de cada zona y punto específico del aeropuerto y su sistema BHS, según el mapa. Color, para identificar el tipo de ruta de los vuelos (doméstica, internacional). Tamaño para el número de maletas que salieron Forma, para representar las maletas (aunque no presta una función para una tarea específica, se usa para diferenciar los puntos usados previamente en la representación de otros atributos y de los rombos usados para las aeronaves). <p>Decisiones de diseño:</p> <ul style="list-style-type: none"> Se utiliza el scatterplot sobre el mapa (simulado) de un sistema BHS y el recorrido que tienen que realizar las maletas. Como se conocen las zonas más importantes de ese recorrido en relación con el proceso de equipajes, esta representación sirve para dar una idea de lo que se puede obtener al contar con un mapa del sistema real, así como sus datos de tracking. El uso del mapa permite aprovechar las características preexistentes del usuario en relación con su conocimiento de la ubicación de cada punto, considerando las implicaciones que puede tener la congestión de un punto del recorrido en un momento dado respecto del entorno en esa zona del aeropuerto. En el caso de las make-up areas, se usa la cantidad (11) y clasificación del aeropuerto El Dorado según (OPAIN-S.A., 2019), pág. 114. Se incluye el tipo de ruta considerando la distribución de las áreas de make-up en el mapa (parte inferior): existen unas make-up areas para vuelos internacionales y otras para vuelos domésticos. Esto permite identificar casos de asignación incorrecta de áreas a vuelos. 	T12

Representación						Codificación	Tarea																																																																																																																																
						<ul style="list-style-type: none"> El tamaño de los puntos con base en las maletas permite conocer si en un momento del día hubo asignación en exceso a un área específica. La distribución en el mapa también permite conocer si hubo áreas sin asignación, o potencialmente fuera de servicio. 																																																																																																																																	
<p>Cifras por hora, destino y vuelo</p> <table border="1"> <thead> <tr> <th>Hora de Salida</th> <th>Casos</th> <th>IndMB</th> <th>Vuelos</th> <th>Pax</th> <th>Maletas</th> </tr> </thead> <tbody> <tr><td>00</td><td>42</td><td>40,46</td><td>10</td><td>1038</td><td>731</td></tr> <tr><td>01</td><td>53</td><td>34,69</td><td>13</td><td>1528</td><td>1427</td></tr> <tr><td>02</td><td>77</td><td>27,17</td><td>18</td><td>2834</td><td>3037</td></tr> <tr><td>03</td><td>11</td><td>18,61</td><td>5</td><td>591</td><td>529</td></tr> <tr><td>04</td><td>16</td><td>16,84</td><td>5</td><td>950</td><td>765</td></tr> <tr><td>06</td><td>2</td><td>7,94</td><td>1</td><td>252</td><td>323</td></tr> <tr><td>10</td><td>16</td><td>23,39</td><td>6</td><td>684</td><td>490</td></tr> <tr><td>11</td><td>25</td><td>32,98</td><td>7</td><td>758</td><td>629</td></tr> <tr><td>12</td><td>23</td><td>14,73</td><td>12</td><td>1561</td><td>1507</td></tr> <tr><td>13</td><td>10</td><td>11,72</td><td>7</td><td>853</td><td>716</td></tr> <tr><td>14</td><td>19</td><td>16,77</td><td>8</td><td>1133</td><td>1021</td></tr> <tr><td>15</td><td>7</td><td>9,55</td><td>5</td><td>733</td><td>517</td></tr> <tr><td>16</td><td>2</td><td>9,17</td><td>2</td><td>218</td><td>128</td></tr> <tr><td>17</td><td>25</td><td>14,16</td><td>11</td><td>1765</td><td>1701</td></tr> <tr><td>18</td><td>38</td><td>21,30</td><td>13</td><td>1784</td><td>1715</td></tr> <tr><td>19</td><td>12</td><td>16,64</td><td>6</td><td>721</td><td>718</td></tr> <tr><td>20</td><td>21</td><td>17,77</td><td>10</td><td>1182</td><td>919</td></tr> <tr><td>21</td><td>11</td><td>19,10</td><td>5</td><td>576</td><td>330</td></tr> <tr><td>22</td><td>4</td><td>11,98</td><td>3</td><td>334</td><td>36</td></tr> <tr><td>23</td><td>28</td><td>20,63</td><td>10</td><td>1357</td><td>804</td></tr> <tr> <td>Total</td><td>442</td><td>21,20</td><td>157</td><td>20852</td><td>18043</td><td></td><td></td></tr> </tbody> </table>	Hora de Salida	Casos	IndMB	Vuelos	Pax	Maletas	00	42	40,46	10	1038	731	01	53	34,69	13	1528	1427	02	77	27,17	18	2834	3037	03	11	18,61	5	591	529	04	16	16,84	5	950	765	06	2	7,94	1	252	323	10	16	23,39	6	684	490	11	25	32,98	7	758	629	12	23	14,73	12	1561	1507	13	10	11,72	7	853	716	14	19	16,77	8	1133	1021	15	7	9,55	5	733	517	16	2	9,17	2	218	128	17	25	14,16	11	1765	1701	18	38	21,30	13	1784	1715	19	12	16,64	6	721	718	20	21	17,77	10	1182	919	21	11	19,10	5	576	330	22	4	11,98	3	334	36	23	28	20,63	10	1357	804	Total	442	21,20	157	20852	18043			
Hora de Salida	Casos	IndMB	Vuelos	Pax	Maletas																																																																																																																																		
00	42	40,46	10	1038	731																																																																																																																																		
01	53	34,69	13	1528	1427																																																																																																																																		
02	77	27,17	18	2834	3037																																																																																																																																		
03	11	18,61	5	591	529																																																																																																																																		
04	16	16,84	5	950	765																																																																																																																																		
06	2	7,94	1	252	323																																																																																																																																		
10	16	23,39	6	684	490																																																																																																																																		
11	25	32,98	7	758	629																																																																																																																																		
12	23	14,73	12	1561	1507																																																																																																																																		
13	10	11,72	7	853	716																																																																																																																																		
14	19	16,77	8	1133	1021																																																																																																																																		
15	7	9,55	5	733	517																																																																																																																																		
16	2	9,17	2	218	128																																																																																																																																		
17	25	14,16	11	1765	1701																																																																																																																																		
18	38	21,30	13	1784	1715																																																																																																																																		
19	12	16,64	6	721	718																																																																																																																																		
20	21	17,77	10	1182	919																																																																																																																																		
21	11	19,10	5	576	330																																																																																																																																		
22	4	11,98	3	334	36																																																																																																																																		
23	28	20,63	10	1357	804																																																																																																																																		
Total	442	21,20	157	20852	18043																																																																																																																																		



Prototipo de Herramienta de analítica visual para el manejo de irregularidades de equipajes

Guía para el Usuario

Menú

Página anterior → Restablecer: Elimina los filtros y retorna la vista a su configuración inicial

← Página siguiente

Resumen, Vuelos y Demoras, Clima, Contexto (BOG), Entradas (BOG), Salidas (BOG), Instrucciones

Cada una de las vistas de análisis disponibles

Guía de uso (usted está aquí ⓘ)

Acciones principales

Selección de ítems de datos para obtener filtros sobre las demás visualizaciones.

Cifras por Región

EUROPA: 12 \$50 mil. NORTE AMÉRICA: 13 \$53 mil. SUR AMÉRICA: 14 \$1 mil. SUD AMÉRICA: 15 \$1.5 mil.

Mostrar como tabla, Incluir, Excluir, Obtener detalles

También puede excluir un ítem de datos (un valor atípico, por ejemplo). El valor seguirá presente en las demás gráficas. Para volver a incluirlo use Restablecer.

Al seleccionar un ítem sobre ciertas gráficas, se activa un botón para usar este filtro en la siguiente vista de detalle.

Puede obtener la misma navegación dando click derecho sobre el ítem, opción 'Obtener detalles', y seleccione las vistas disponibles para usar el ítem como filtro.

Combinación de filtros

Para combinar filtros, seleccione un ítem de datos y luego use la tecla **Ctrl** para seleccionar otro ítem, ya sea en la misma gráfica o en una diferente.

Por ejemplo, para ver la tendencia de daños en BOG, seleccione el ítem BOG...

Distribución de casos y MB por periodo

Top aeropuertos (casos): BOG: 4319, LIM: 718

Distribución casos y costos por periodo

Cifras por Tipo de irregularidad

... y luego, con **Ctrl** sostenido, seleccione el ítem "daños". Verá los filtros aplicados sobre la gráfica de tendencia:

Al seleccionar esta opción, ya puede seleccionar el ítem de datos de la gráfica para ver sus detalles en el siguiente nivel. Para interactuar con la nueva vista y usarla como filtro en otras gráficas, desactive esta opción.

Use estas opciones para ampliar la gráfica a pantalla completa, o para acceder a otras opciones del menú.

Para ver solo un mes, seleccione el nombre (enero). Para seleccionar una categoría dentro del mes, seleccione la parte correspondiente en la barra (la sección azul para "demoras", por ejemplo).

Vistas con jerarquías

Estos controles indican que el objeto visual maneja jerarquías

Distribución de casos y MB

Tipo: Daños, Demoras, Pérdidas, Saqueos, IndMB

Activar exploración en profundidad (siguiente nivel)

Subir de nivel

Ir al siguiente nivel (solo muestra el siguiente nivel y agrega los valores)

Explande al siguiente nivel (muestra los dos niveles)

Al seleccionar esta opción, ya puede seleccionar el ítem de datos de la gráfica para ver sus detalles en el siguiente nivel. Para interactuar con la nueva vista y usarla como filtro en otras gráficas, desactive esta opción.

El título de los objetos con jerarquía indica los niveles que se pueden explorar

Casos por Segmento y Vuelo

Casos por Tipo de Causa y Cód.Causa (RU)

Casos por Tipo de Causa y Cód.Causa (RU)

Segmento: BOG-MOE, BOG-CLO, BOG-CTG, BOG-BAQ, BOG-BGA. Causa: daños, errores, faltas, aerop., incumpl.

área de trabajo

Elaborado por:

Emilio Alvarado
ea.alvarado67@uniandes.edu.co
@eaalvarados

Última actualización: Mayo 2020
v1.11

Disclaimer
Los datos de irregularidades en este prototipo son de prueba, basados en el comportamiento real de la industria. Los datos de vuelos asociados con las irregularidades son basados en rutas reales. Se conservan los códigos usados por la IATA para aeropuertos, aeronaves, países, entre otros. El indicador MB no debe tomarse como dato oficial de ninguna aerolínea.
Se utilizan otros datos de contexto de entidades oficiales (Aerocivil, EIDrado), así como estadísticas de terceras partes (FlightStats) para un día específico.

Figura 43. Guía para el usuario

Consideraciones adicionales

Durante el diseño de la herramienta de visualización se evaluaron diferentes alternativas de modismos para resolver ciertas tareas según el conjunto de datos y tipos de atributos. Algunos de los modismos evaluados en la fase de diseño se listan en la Tabla 49 del **Anexo A**. Muchos de ellos no pudieron ser incorporados en la solución por uno o varios de los siguientes factores:

- Escalabilidad: por el volumen de datos implicaban una alta densidad de ítems en el gráfico, que hacía imperceptible la información relevante.
- Desempeño: al coordinarlos con otras vistas se degradaba considerablemente el desempeño, también asociado al volumen de datos.
- Costo: implicaba la adquisición de un paquete de licenciamiento.
- Uso de componentes adicionales: algunos requerían el uso de R o Python.
- Restricción de uso: algunos que los que requerían componentes adicionales, funcionaban en entorno de diseño pero no en la publicación sin costo a PowerBi Service.

4.2.5. Evaluación

Durante la etapa de validación de la aproximación analítica utilizada, se identificaron algunos hallazgos relevantes, los cuales se describen a continuación.

4.2.5.1. *Hallazgos relevantes en datos*

En relación con los datos de demoras en vuelos, se identifica que en términos generales que los niveles de demora mayores a 45 minutos generan un indicador MB alto. No obstante, no siempre es mayor que aquél para los niveles de demora menores a 15 minutos, por ser éste el nivel en que se ubica la mayoría de vuelos. Adicionalmente, los vuelos sin demora están asociados a la mayoría de las irregularidades presentadas. No se puede asociar un vuelo sin demoras a no ocurrencia de irregularidades.

Con respecto al clima, hay correspondencia entre las reglas de vuelo, los tipos de nubosidad y la visibilidad. Sin embargo, es difícil encontrar casos en los cuales las demoras en los vuelos sean ocasionadas por condiciones climática adversas, y, por tanto, no es posible asociar las irregularidades presentadas en los vuelos con demora a las condiciones

climáticas. Aunque existen algunos códigos de causa (RL) indirectamente relacionados con el clima (62 y 65), no se encuentran casos en donde se atribuyeron como causa. Sería interesante poder indagar con mayor detalle este tipo de casos con los especialistas.

En relación con los datos de contexto, el volumen de vuelos realizados y de pasajeros transportados en los datos del caso de estudio son congruentes con las irregularidades presentadas. Sin embargo, no guardan una relación con los vuelos y pasajeros transportados por las demás aerolíneas según los datos del aeropuerto El Dorado en el periodo de tiempo analizado.

La vista de contexto para los vuelos que ingresan al aeropuerto El Dorado (*Entradas BOG*) permite entender la operación en zonas de entrega de equipaje y puertas de ingreso de pasajeros en un momento específico del día. Es posible identificar carruseles de equipaje con sobreasignación y otros sin uso en momentos específicos del día. También es posible observar pocos casos en los cuales más de un vuelo se presentan en la misma puerta de ingreso en un momento dato. En este punto es importante anotar que aunque la granularidad de los datos de tiempo se manejó a nivel de horas, y que la operación puede ser mucho más dinámica, la propuesta de análisis cumple con la tarea de identificar este tipo de comportamientos, que pueden ser extrapolados a un entorno real con los datos correspondientes y el refinamiento del detalle de tiempo que se desea incorporar.

Lo mismo sucede en el caso de la vista de contexto para los vuelos que salen del aeropuerto (*Salidas BOG*). Aunque la disposición del sistema BHS es simulada, así como la asignación de los vuelos según su ruta a las áreas de *make-up*, es posible identificar fácilmente comportamientos en esta distribución en puntos específicos en un momento del día, como áreas sobrecargadas y otras sin asignación.

4.2.5.2. *Hallazgos en el uso de modismos de visualización*

El uso del modismo *scatterplot* sobre mapas que representan la distribución y topología de las zonas del aeropuerto es importante en la identificación de comportamientos a partir de los datos. Tal es el caso de la asignación de carruseles de equipaje o de *make-up areas* a rutas domésticas o internacionales en una zona diferente. Aunque esto podría tener causas justificadas según las reglas de operación del aeropuerto o sus planes de contingencia, es interesante poder ver ese tipo de situaciones reflejadas por los datos.

4.2.6. Despliegue

El despliegue de la solución en esta fase sigue el diseño presentado en la vista física de la Figura 36. No hay consideraciones adicionales, aunque se experimentan tiempos más altos en la publicación hacia PowerBi Service, dato el nuevo volumen de datos.

4.2.7. Retroalimentación

4.2.7.1. Pruebas funcionales

En esta etapa se consolidan las pruebas realizadas con usuarios tanto de la fase descriptiva como la de análisis diagnóstico.

Se diseñó un conjunto de pruebas con una pregunta de análisis de complejidad baja, una de media y una de alta para cada vista de la solución (la cual se identificó como *sección de pruebas*), para un total de 18 preguntas. Las preguntas de complejidad baja atendían a realizar una observación sin interacción sobre la vista (*overview*). Las preguntas de complejidad media requerían de una acción de interacción sobre alguno de los objetos de la vista (*zoom & filter*), generalmente filtros, y luego una acción de observación. Las preguntas de complejidad alta requerían al menos dos acciones de interacción sobre la vista (*details on demand*), generalmente filtros combinados y en pocos casos *drill-down* sobre objetos visuales con jerarquías. Se incluyeron preguntas de selección múltiple (generalmente para complejidad baja y media) y de respuesta abierta (para complejidades media y alta). Las pruebas fueron aplicadas a una población de 22 usuarios.

Los resultados de las pruebas se resumen en la Figura 44. En general, el porcentaje de aciertos fue mayor para las pruebas de complejidad baja, menor para las de complejidad media y el más bajo se obtuvo en las preguntas de complejidad alta. Las secciones con mejor porcentaje de aciertos fueron la vista de contexto y la de vuelos y demoras, y los resultados más bajos se obtuvieron en la vista de resumen y la de contexto-salidas.

En términos generales las preguntas de complejidad alta y media obtuvieron los mejores porcentajes de aciertos en todas las vistas, a excepción de la vista de contexto y la de salidas, en esta última todas las preguntas tuvieron el mismo nivel de acierto.

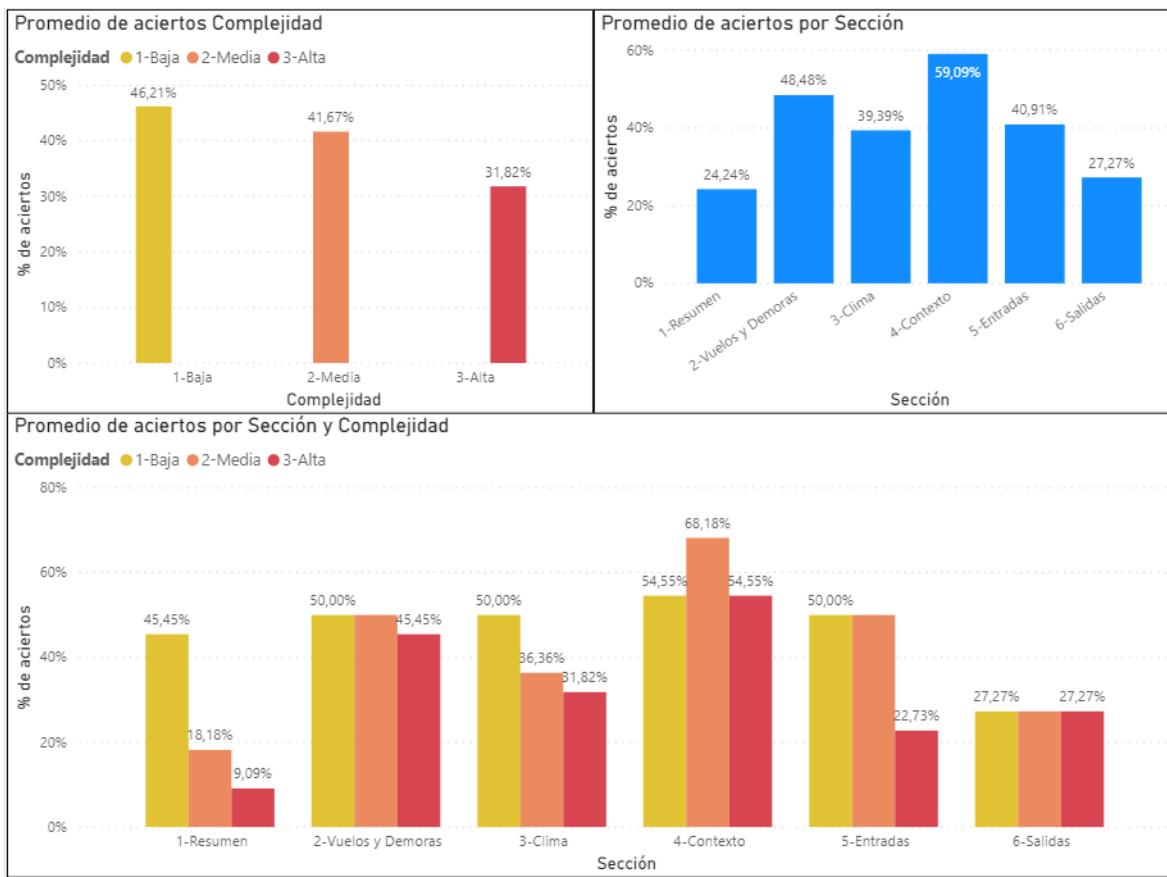


Figura 44. Resultados de las pruebas con usuarios

Las preguntas de respuesta abierta permiten identificar cierta dificultad en la vista de salidas, lo cual es posible atribuir a la ausencia de conocimientos previos en el dominio de negocio. El bajo porcentaje en la vista de resumen se atribuye a la confusión que tuvieron los usuarios en el manejo de los indicadores, especialmente en los diagramas de *scatterplot*, que manejaban tanto el indicador MB como los costos, y las preguntas hacían alusión a ambos. El alto porcentaje en la pregunta de complejidad baja de la vista de contexto se puede explicar por la ausencia de elementos de interacción, dadas las restricciones en la coordinación de vistas de esta sección por la agregación de los datos a distintos niveles entre la información de la Aerocivil y El Dorado, se optó por otra pregunta de observación sin interacción, aunque con mayor complejidad que la de complejidad baja.

La vista de vuelos y demoras es la segunda con mejor porcentaje de aciertos. Se identifica que en esta vista, a pesar de las preguntas que exigían varias tareas de interacción involucrando filtros, cuenta en la parte superior con filtros tradicionales que permitían al usuario contar con la información requerida sin tener que combinar acciones sobre las

gráficas. Este punto es importante para incorporar mejoras en vistas con complejidad alta en términos de objetos visuales, cantidad de variables o tareas a realizar.

La vista de salidas presentó los niveles más bajos en las respuestas de complejidad baja y media. Las preguntas utilizadas aquí fueron todas abiertas, las cuales invitaban a realizar observaciones y analizar la información presentada. En muchos de los casos se obtuvo respuestas demasiado cortas o sin información, que se pueden asociar a la longitud de la prueba, al ser la última sección, más que a la dificultad de las tareas o el uso de la herramienta.

Las tareas que implicaron una o varias interacciones y observaciones fueron las de más bajo desempeño. Se identifica una dificultad generalizada para efectuar tareas que impliquen interacción por medio de combinación de filtros. Esto permite anticipar la necesidad de preparar material y realizar actividades de entrenamiento a los usuarios, con el fin de adquirir destrezas en el uso de las interacciones. Aunque se identificaron varios casos en donde se evidenció la omisión de las instrucciones, también es posible incorporar mejoras en temas de usabilidad, como colores que indiquen cuando una situación o un indicador asume un comportamiento anormal, así como hacer más explícita la manera en que se pueden eliminar y combinar los filtros, o restablecer la configuración inicial.

4.2.7.2. Pruebas de desempeño

Se realizaron las mediciones en entorno de desarrollo, en las mismas condiciones descritas para la fase de análisis descriptivo en esta mista etapa. Los indicadores se resumen en la Tabla 48.

En términos generales, estas vistas incorporan mayor número de variables y conjuntos de datos que los utilizados en la fase de análisis descriptivo. Los tiempos de respuesta son mayores en la gran mayoría de casos. Las vistas que reúnen gran cantidad de objetos visuales y manejan coordinación entre los mismos presentan tiempos de respuesta más altos, tal es el caso de la vista de Vuelos y Demoras y la vista de Clima.

La vista de Contexto, a pesar de manejar conjunto de datos de menor tamaño por ser cifras agregadas de entidades externas, presenta tiempos muy altos respecto de otras vistas con mayor complejidad. Esto puede atribuirse a la necesidad de unir estos conjuntos de datos con el de irregularidades para presentar el número de casos, lo cual permite identificar oportunidades de mejora en la construcción de las consultas internas (*queries, joins*).

Tabla 48. Mediciones de tiempo en las interacciones para la fase diagnóstica

Vista	Medición	No. de objetos visuales con interacción / coordinación
Vuelos y Demoras	<ul style="list-style-type: none"> Cross-highlighted + Distribución de Maletas y Pax 2174 + Vuelos por # de conexiones* y nivel de demora 2170 + Tipo de Ruta y Aeronave 2168 + Slicer 2154 + Slicer 2151 + Vuelos por nivel de demora 2149 + Indicador MB 2148 	11
Clima	<ul style="list-style-type: none"> Cross-highlighted + Demora promedio (mins) y casos por periodo 1673 + Indicador MB 1664 + Reglas de Vuelo (destino) 1657 + Visibilidad promedio en estación destino por hora del día 1657 + Demora promedio en el origen por hora del día 1654 + Visibilidad promedio en estación origen por hora del día 1647 + Reglas de Vuelo (origen) 1647 + Demora promedio en el destino por hora del día 1644 + Responsable por # de casos 1641 + Nubosidad en origen 1638 + Nubosidad en destino 1635 	12
Contexto	<ul style="list-style-type: none"> (+) Drilled down/up + Indicador MB 2996 + EIDorado - Estadísticas diarias - Vuelos 2869 + Card 2702 + EIDorado - Estadísticas diarias - Pasajeros 2629 + Aerocivil (mensual) - Vuelos entrando a BOG 2499 + Card 2361 + Aerocivil (mensual) - Vuelos por Empresa 2282 	5
Contexto – Entradas BOG	<ul style="list-style-type: none"> Cross-highlighted + Hora de Llegada 2300 + Indicador MB del Día 2225 + Top Vuelos por Aerolínea 2145 + Casos por Hora de Llegada (Informativo) 2076 + Cifras por Origen y Vuelo 2069 + Card 1792 + Vuelos y Pasajeros* por Puerta de Ingreso 1643 	5

Vista	Medición	No. de objetos visuales con interacción / coordinación																																
Contexto – Salidas BOG	<table border="1"> <tr><td>Cross-highlighted</td><td>-</td></tr> <tr><td> Cifras por hora, destino y vuelo</td><td>4007</td></tr> <tr><td> Casos por Tipo de Ruta e Irregularidad</td><td>3938</td></tr> <tr><td> Casos por Nivel de Demora y Tipo de Irreg.</td><td>3789</td></tr> <tr><td> Indicador MB</td><td>3316</td></tr> <tr><td> Card</td><td>3157</td></tr> <tr><td> Card</td><td>2960</td></tr> <tr><td> Card</td><td>2607</td></tr> <tr><td> Card</td><td>2438</td></tr> </table> <table border="1"> <tr><td>Cross-highlighted</td><td>-</td></tr> <tr><td> Casos por Nivel de Demora y Tipo de Irreg.</td><td>906</td></tr> <tr><td> Casos por Tipo de Ruta e Irregularidad</td><td>906</td></tr> <tr><td> Cifras por hora, destino y vuelo</td><td>901</td></tr> <tr><td> Card</td><td>896</td></tr> <tr><td> Card</td><td>894</td></tr> <tr><td> Card</td><td>891</td></tr> </table>	Cross-highlighted	-	Cifras por hora, destino y vuelo	4007	Casos por Tipo de Ruta e Irregularidad	3938	Casos por Nivel de Demora y Tipo de Irreg.	3789	Indicador MB	3316	Card	3157	Card	2960	Card	2607	Card	2438	Cross-highlighted	-	Casos por Nivel de Demora y Tipo de Irreg.	906	Casos por Tipo de Ruta e Irregularidad	906	Cifras por hora, destino y vuelo	901	Card	896	Card	894	Card	891	4
Cross-highlighted	-																																	
Cifras por hora, destino y vuelo	4007																																	
Casos por Tipo de Ruta e Irregularidad	3938																																	
Casos por Nivel de Demora y Tipo de Irreg.	3789																																	
Indicador MB	3316																																	
Card	3157																																	
Card	2960																																	
Card	2607																																	
Card	2438																																	
Cross-highlighted	-																																	
Casos por Nivel de Demora y Tipo de Irreg.	906																																	
Casos por Tipo de Ruta e Irregularidad	906																																	
Cifras por hora, destino y vuelo	901																																	
Card	896																																	
Card	894																																	
Card	891																																	

También se observa que las vistas que contienen matrices de datos generan tiempos de respuesta más altos en la actualización requerida tras una interacción, así como las acciones de *drill-down* sobre ellas cuando manejan jerarquías. De otro lado, la situación en la que los tiempos de respuesta más altos se dan en los primeros usos de la solución, se evidencia en mayor medida en la vista de contexto-salidas BOG. Los tiempos de respuesta son de los más altos, sin embargo, tras el uso, se obtienen tiempos entre los mejores de toda la solución. Es importante analizar esta situación con mayor profundidad en términos del número de ítems que se manejan, los conjuntos de datos involucrados, las operaciones de agregación requeridas y también el uso de recursos de máquina durante esos dos momentos específicos.

4.3. Análisis y discusión

El proceso de gestión de equipajes presenta diferentes retos en la búsqueda de reducir el número de irregularidades presentadas y por tanto el indicador MB, uno de los más relevantes a la hora de comparar los niveles de servicio entre compañías aéreas. Algunos de esos retos están asociados a condiciones fuera del control de las aerolíneas, como el manejo de conexiones de vuelos con otras compañías, las restricciones de peso o espacio de las aeronaves, las fallas en los sistemas de transporte de equipajes (BHS) dentro de los aeropuertos, las revisiones de equipaje según políticas de seguridad aeroportuaria, o la manipulación del equipaje por parte del personal de operación propio de los aeropuertos o

de las empresas en concesión. Otras están relacionadas con factores sobre los cuales la compañía puede influir, como el etiquetado correcto de las maletas, la coordinación entre el embarque o desembarque de pasajeros y sus maletas o la correcta selección y clasificación del equipaje hacia los vuelos correspondientes.

También es importante recordar que este proceso se deriva de y está estrechamente ligado al transporte de pasajeros, con lo cual el estudio de los factores alrededor de la operación de vuelos es de crucial importancia para determinar elementos que influyan indirectamente en las irregularidades de equipajes. Áreas de análisis como las demoras presentadas en los vuelos, tanto en el origen como en el destino, son de gran importancia considerando que el mayor tipo de irregularidad de equipajes que se presenta es el de demora. Esto implica, a su vez, el análisis de otros factores relacionados con las demoras en los vuelos, como las condiciones climáticas y la operación aeroportuaria.

En relación con otro tipo de irregularidades, como los daños, las pérdidas o los saqueos, es importante incorporar también el análisis del proceso de gestión de operaciones terrestres, incluyendo la gestión de personal, recursos, e infraestructura aeroportuaria, entre otros. A lo anterior es necesario agregar que una de las causas más recurrentes asociadas con las irregularidades de tipo demora tiene que ver con problemas en la infraestructura aeroportuaria.

La diversidad de actores y compañías que participan en estos procesos constituye un reto adicional en la gestión de equipajes, especialmente por la distribución de las responsabilidades, la variedad y heterogeneidad de los sistemas de información y la distribución, falta de disponibilidad y acceso a los datos que el proceso de gestión de equipajes y los demás procesos relacionados producen en todas sus etapas.

Durante el proceso de gestión de irregularidades se realizan tareas de investigación que finalizan en la entrega del equipaje al pasajero o en su defecto en el pago de una compensación. Con respecto a los datos, se asigna un aeropuerto responsable al caso (incrementando el valor del indicador MB en ese sitio) y un código de causa (RL). Estas causas hacen parte de una codificación estandarizada que abstrae la complejidad de este proceso y la reduce a 48 posibles razones. Como ya se mencionó, una de las causas más frecuentes en los datos es ‘problemas en la infraestructura aeroportuaria’, sin embargo esto puede obedecer a muchas variables dentro de dicha infraestructura y ese detalle no aparece reflejado en la información. Por consiguiente, incluir información adicional, como la

infraestructura aeroportuaria y el comportamiento de la operación, a las variables de análisis y correlacionarla con las irregularidades dentro del alcance del análisis diagnóstico, representa una oportunidad para identificar posibles razones de causa más detalladas que habiliten la elaboración de planes de acción sobre elementos más concretos.

El planteamiento del proyecto teniendo como objetivo el diseño de una solución que apoye el análisis de irregularidades en el manejo de equipajes, constituye una propuesta para generar valor a partir de la información. Este tipo de iniciativas maximizan sus beneficios potenciales en la medida en que las fuentes de información identificadas cuenten con la disponibilidad y acceso, dentro de las políticas de seguridad y manejo de la información de sus propietarios.

En el caso de las actividades de generación de datos, durante el proceso de simulación para el análisis de contexto respecto a las zonas del aeropuerto El Dorado (especialmente en las salidas), estas implicaron un proceso dispendioso y con un número bastante alto de reglas de negocio que es necesario considerar para tener un acercamiento mayor a las condiciones reales. Sin embargo, el ejercicio permite identificar el potencial y valor que este tipo de actividad genera ante la ausencia de datos.

Con relación a la implementación y uso de la solución, el análisis descriptivo apoyado en visualizaciones permite a los especialistas del negocio la exploración de los datos de una manera más flexible y acorde a sus necesidades e intereses. Lo mismo sucede en la fase de análisis diagnóstico, donde el espectro del análisis se incrementa en la medida en que se incorporan nuevas fuentes de información, y así mismo las necesidades de exploración. En esta parte también se incrementan los esfuerzos en el diseño de visualizaciones y se hace más necesaria la identificación de áreas de enfoque para el análisis. En tal sentido, cobra relevancia la combinación de este enfoque con los análisis asistidos (o automatizados).

Uno de los casos que refleja esta necesidad se presentó en la vista de condiciones climáticas, en donde fue evidente la dificultad para encontrar asociaciones entre las demoras de los vuelos, el clima y las irregularidades. Para situaciones en donde confluyen varios conjuntos de datos, así como mayores volúmenes, el análisis asistido es una buena alternativa para buscar rápidamente patrones y correlaciones entre los datos. No obstante demanda esfuerzos en el modelamiento y sigue necesitando de la retroalimentación del especialista de negocio. De allí el beneficio de utilizar las dos aproximaciones.

Respecto al diseño de las visualizaciones, el uso del marco de trabajo mencionado en la metodología facilita el enfoque en el usuario a través de la identificación de tareas alineadas con sus objetivos de análisis.

Para la implementación de la solución sobre un producto como PowerBi, es importante resaltar que existe una amplia oferta de objetos visuales personalizados que pueden ampliar la oferta de modismos de visualización a utilizar durante las decisiones de diseño. También existe la opción de construir objetos visuales personalizados utilizando herramientas como R, Python o D3. Este proyecto aporta una contribución en la catalogación de modismos no convencionales construidos para PowerBi (presentes en el Anexo A), de tal manera que puedan ser evaluados como alternativa a la hora de dar continuidad a este proyecto o iniciar otros donde se utilice la misma plataforma.

En términos del desempeño, si bien el comportamiento general es adecuado, aquellas secciones con mayor número de objetos interactuando tienen los tiempos más altos. Existe la posibilidad de refinar la propuesta no solo reduciendo el número de gráficos sino seleccionando en qué casos la coordinación o el resaltado entre ellos es necesario, y deshabilitando toda interacción con los objetos que no despliegan información para interacción. Adicionalmente, es posible refinar la arquitectura interna en la capa de almacenamiento a través del uso de una base de datos, que permita a la solución escalar en volumen con el paso del tiempo. También es posible mejorar la mantenibilidad de la capa de procesamiento a través de la incorporación de *pipelines* que ejecuten las tareas de integración, limpieza, calidad y generación de datos sobre cada nueva fuente que se incorpore. Esto mismo aplica para la generación de modelos en el análisis asistido o en análisis más avanzados como el predictivo.

Los resultados de las pruebas funcionales permiten identificar la necesidad de entrenamiento previo a los usuarios, además de la elaboración de una guía más dinámica. Algunas de las acciones de interacción con las visualizaciones no son tan intuitivas como se espera, y la combinación de varias de estas acciones eleva el nivel de dificultad para la ejecución de las tareas. Este es el caso de los filtros combinados a través de las gráficas y las acciones de *drill-down* en objetos con jerarquías. Se identifican también aspectos a mejorar que están relacionados con el número elevado de gráficas en una misma vista y la consistencia en el uso de filtros tradicionales a lo largo de las diferentes secciones.

De otro lado, es clave obtener la retroalimentación del usuario en cada iteración, de tal manera que se puedan incorporar mejoras continuamente. En este proyecto, dadas las restricciones de tiempo, fue necesario realizar un solo ciclo de pruebas que reunió todas las fases.

Capítulo 5. Conclusiones y recomendaciones

5.1. Conclusiones

Dentro de las problemáticas y necesidades más comunes en los procesos de gestión de equipajes e irregularidades se encuentran el alto número de variables que afectan la operación de los vuelos, la diversidad de actores y sistemas involucrados, la distribución de la información y la carencia de iniciativas encaminadas a sacar el mayor provecho de los datos que producen los procesos en cada una de sus etapas. Adicionalmente, la clasificación de las causas de las irregularidades no refleja la complejidad real de estos procesos.

El uso de ASUM-DM como metodología, combinada con procesos y marcos de trabajo de analítica visual (Keim, Munzner), permite establecer una estructura de abordaje de proyectos basados en información, que habilita el diseño de soluciones centradas en el usuario y sus objetivos de análisis, para contribuir a las problemáticas identificadas en el manejo de irregularidades de equipajes. Sumado a lo anterior, la aplicación de técnicas de análisis asistido de manera conjunta con el diseño y uso de visualizaciones genera el mayor valor en la identificación de patrones o correlaciones entre los datos, y de esta manera contribuye a la segmentación de áreas específicas de análisis para determinar el enfoque de las siguientes etapas dentro de un proceso iterativo.

El diseño de la herramienta a partir de los datos de irregularidades, vuelos, demoras, condiciones climáticas, contexto e infraestructura del aeropuerto, así como la simulación de datos para representar el comportamiento de entradas y salidas de vuelos y equipajes; permitió reunir en un solo punto diferentes fuentes de datos y habilitar capacidades de análisis descriptivo de las irregularidades presentadas y diagnóstico sobre las demás variables relacionadas.

Por último, es importante resaltar que la información disponible en los sistemas actuales, independiente de su propietario, puede ser utilizada de una mejor manera si se articulan estrategias de mutuo beneficio entre aeropuertos y aerolíneas, que apoyen la generación de proyectos, la adquisición de productos o la construcción de soluciones que permitan explotar la información producida por todos los procesos y áreas, para su correspondiente optimización y para apoyar la toma de decisiones basadas en datos en todos los niveles (estratégico, táctico y operativo).

5.2. Recomendaciones y trabajo futuro

Dado el enfoque del trabajo en la analítica visual, los análisis asistidos estuvieron acotados a la correlación de atributos cuantitativos. Es posible profundizar en el análisis de correlación de atributos categóricos, utilizando técnicas estadísticas conocidas, y luego establecer tareas para el diseño de visualizaciones que permitan comparar los dos resultados.

De otro lado, el acceso a una gran cantidad de datos presentes en los sistemas BHS y BRS de los aeropuertos permitiría contar con un insumo invaluable para el establecimiento de iniciativas similares. En ese mismo sentido, se abriría la puerta para estudios adicionales, como los proyectos de análisis de información espacio-temporal y en movimiento, así como ejercicios de simulación mucho más precisos y formales.

Profundizar en el análisis diagnóstico planteado en este trabajo, así como en la identificación de variables que inciden en las irregularidades, daría lugar a la generación de planes de acción mucho más concretos en la mitigación del riesgo de ocurrencia de irregularidades de equipajes. Esto complementaría muy bien los resultados de una posible fase adicional enfocada en analítica predictiva.

En términos de diseño e implementación de visualizaciones, es posible evaluar modismos no convencionales para enfrentar tareas y conjuntos de datos complejos, así como diseñar e implementar objetos visuales personalizados basados en otros lenguajes y frameworks, como D3, Python o R; y que funcionen adecuadamente en el entorno de PowerBi, en términos de coordinación con otros objetos, escalabilidad y adecuado desempeño.

Anexo A. Listado de modismos de visualización para PowerBi

Como parte del trabajo de investigación y diseño de la herramienta de visualización, se exploraron modismos no convencionales en Microsoft AppSource. Algunos de ellos se enumeran en la Tabla 49 el según el propósito o tipo de datos que maneja. Este listado puede ser de utilidad para la incorporación de mejoras en la evolución del trabajo, o para abordar nuevos proyectos con conjuntos de datos diferentes. Es importante consultar si cada alternativa se encuentra certificada y vigente en el AppSource, y evaluarla en términos de la escalabilidad y desempeño para cada proyecto.

Tabla 49. Modismos de visualización no convencionales para PowerBi

Tarea/Tipo de datos	Nombre	Ubicación
Distribución	Sunburst	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104380767
	Waffle Chart	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104381049?tab=Overview
	RadarChart	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104380771
	ClusterMap	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104380806
	Mekko Chart	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104380785
	Table Heatmap	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104380818
	Funnel plot	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104381066
	Stream graph	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104380772
	Heat Streams	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104381183?tab=Overview
	Variance Chart	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA200000182?tab=Overview
Datos temporales	Bullet Chart	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104380755?tab=Overview
	Timeline	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104381377?tab=Overview
	Calendar	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104381146?tab=Overview
Monthly-Calendar	Monthly-Calendar	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104381096?tab=Overview

Tarea/Tipo de datos	Nombre	Ubicación
	Timeline Storyteller	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104381136?tab=Overview
	Timeseries Forecasting	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104380816
Redes	Advanced Network Visual	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104381419?tab=Overview
	Outside label network	https://zoomcharts.com/en/microsoft-power-bi-custom-visuals/gallery/outside-label-settings-using-advanced-network-visual?download&async=1
	Force-Directed Graph	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104380764
Correlación	Chord	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104380761?tab=Overview
	Decomposition Tree	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104381663?tab=Overview
	Sankey Chart	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104380777
	Correlation plot	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104380814
Mapas	Mapbox Visual	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104381472?tab=Overview
	Drilldown Choropleth	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104381044
	Drilldown Cartogram	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104381045
Acciones personalizadas / Indicadores	Hierarchical Filter	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA200000727?tab=Overview
	HierarchySlicer	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104380820?tab=Overview
	Scroller	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104381018?tab=Overview
	Attribute Slicer	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104380794?tab=Overview
	D3.js visual	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104381354?tab=Overview
	User List	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104381426?tab=Overview
	Pivot Slicer	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104381719?tab=Overview
	Dial Gauge	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104381184?tab=Overview
	Timeline slicer	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104380786?tab=Overview
	Text filter	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104381309
Analítica avanzada	Forecast using Neural Network	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104381428?tab=Overview
	ValQ Modern Digital Planning: (what-if analysis, simulation, scenarios)	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA200000006?tab=Overview
	Decision Tree Chart (requiere R)	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104380817?tab=Overview
	Forecasting with Arima	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104380888
	Association rules	https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA104380815

Bibliografía

- Aerocivil. (2020a). Estadísticas de las actividades aeronáuticas. Retrieved from <http://www.aerocivil.gov.co/atencion/estadisticas-de-las-actividades-aeronauticas>
- Aerocivil. (2020b). Meteorológica aeronáutica para Colombia. Retrieved from http://meteorologia.aerocivil.gov.co/obs_map?aoi=COL
- Al-Serafi, A., & Elragal, A. (2014). *Visual trajectory pattern mining: an exploratory study in baggage handling systems*. Paper presented at the Industrial Conference on Data Mining.
- ALSTEF. (2019a). Airport Baggage Handling Systems. Retrieved from <https://www.airport-suppliers.com/supplier/alstef-automation-s-a/>
- Alstef. (2019b). DCV (Destination Coded Vehicles). Retrieved from <https://www.alstef.com/DCV-Destination-Coded-Vehicles-103>
- Alsyouf, I., Humaid, F., & Al Kamali, S. (2014). *Mishandled baggage problem: causes and improvement suggestions*. Paper presented at the 2014 IEEE International Conference on Industrial Engineering and Engineering Management.
- Amadeus. (2019a). Amadeus Altéa Departure Control - Customer Management. Retrieved from <https://amadeus.com/es/catalogo/aerolineas/amadeus-altea-departure-control-customer-management>
- Amadeus. (2019b). Amadeus Altéa Departure Control - Flight Management. Retrieved from <https://amadeus.com/es/catalogo/aerolineas/amadeus-altea-departure-control-flight-management>
- Amadeus. (2019c). Amadeus Baggage Reconciliation System (BRS). Retrieved from <https://amadeus.com/es/catalogo/aeropuertos/amadeus-baggage-reconciliation-solution-brs>
- Andrienko, N., Andrienko, G., & Gatalsky, P. (2003). Exploratory spatio-temporal visualization: an analytical review. *Journal of Visual Languages & Computing*, 14(6), 503-541.
- AvCodes. (2019). Aviation Codes Central - Aircraft Type Codes. Retrieved from <http://www.avcodes.co.uk/acrtypes.asp>
- Aviation-Edge. (2019). Aviation Edge – Database and API. Retrieved from <https://aviation-edge.com/>
- Chen, W., Huang, Z., Wu, F., Zhu, M., Guan, H., & Maciejewski, R. (2017). VAUD: A visual analysis approach for exploring spatio-temporal urban data. *IEEE Transactions on Visualization and Computer Graphics*, 24(9), 2636-2648.

- Farey. (2014). Reason for Lost Codes. Retrieved from <https://es.scribd.com/doc/201710160/Reason-for-Lost-Codes>
- FlightStats-Cirium. (2019). FlightStats Data Exports. Retrieved from <https://www.flighstats.com/v2/data-exports/home>
- Frey, M. (2014). *Models and Methods for Optimizing Baggage Handling at Airports*. Technische Universität München,
- Frey, M., Kiermaier, F., & Kolisch, R. (2017). Optimizing inbound baggage handling at airports. *Transportation Science*, 51(4), 1210-1225.
- Google. (2020a). Mapa interno del aeropuerto El Dorado - Puertas de abordaje. Retrieved from <https://www.google.com/maps/place/Terminal+2,+Ac.+26+%23103-9,+Bogot%C3%A1/@4.699062,-74.1428911,18z/data=!3m1!5s0x8e3f9cb9fdeb04ab:0xc648fe62f9ccdb5!4m12!1m6!3m5!1s0x8e3f9b646667885b:0xb03f261ee6926ef1!2sAeropuerto+Internacional+El+Dorado!8m2!3d4.6988549!4d-74.1412211!3m4!1s0x8e3f9cb75023bd5b:0xe34ef1658a706ee!8m2!3d4.6986519!4d-74.1417962>
- Google. (2020b). Mapa interno del aeropuerto El Dorado - Zona de reclamo de equipaje. Retrieved from <https://www.google.com/maps/place/Terminal+2,+Ac.+26+%23103-9,+Bogot%C3%A1/@4.6983247,-74.1409983,19z/data=!3m1!5s0x8e3f9cb9fdeb04ab:0xc648fe6e49b19d0!4m12!1m6!3m5!1s0x8e3f9b646667885b:0xb03f261ee6926ef1!2sAeropuerto+Internacional+El+Dorado!8m2!3d4.6988549!4d-74.1412211!3m4!1s0x8e3f9cb75023bd5b:0xe34ef1658a706ee!8m2!3d4.6986519!4d-74.1417962>
- Howson, C., Richardson, J. L., Sallam, R. L., & Kronz, A. (2019). Magic quadrant for analytics and business intelligence platforms. Retrieved Feb, 11, 2019.
- IATA. (2017). IATA Resolution 753/A4A Resolution 30.53 Implementation Guide. In (pp. 85). Online: <https://www.iata.org/whatwedo/ops-infra/baggage/Pages/index.aspx>.
- IATA. (2019a). Baggage Tracking. Retrieved from <https://www.iata.org/whatwedo/ops-infra/baggage/Pages/baggage-tracking.aspx>
- IATA. (2019b). IATA Vision and Mission. Retrieved from <https://www.iata.org/en/about/mission>
- Keim Daniel, K. J., Ellis Geoffrey, Mansmann Florian. (2010). *Solving Problems with Visual Analytics*: Eurographics Association.
- Lufthansa-Systems. (2019). NetLine/Ops ++. Retrieved from <https://www.lhsystems.com/solutions-services/operations-solutions/netlineops>
- Malandri, C., Briccoli, M., Mantecchini, L., & Paganelli, F. (2018). A Discrete Event Simulation Model for Inbound Baggage Handling. *Transportation research procedia*, 35, 295-304.
- Microsoft. (2019a). Customer Service - Microsoft Dynamics 365. Retrieved from <https://dynamics.microsoft.com/es-mx/customer-service/overview/>

- Microsoft. (2019b). PowerBI Interactive Data Visualization BI Tools. Retrieved from <https://powerbi.microsoft.com/en-us/>
- Microsoft. (2020a). PowerBI Visuals in Microsoft AppSource. Retrieved from <https://appsource.microsoft.com/en-us/marketplace/apps?product=power-bi-visuals>
- Microsoft. (2020b). Publish to web from Power BI. Retrieved from <https://docs.microsoft.com/en-us/power-bi/collaborate-share/service-publish-to-web>
- Munzner, T. (2014). *Visualization analysis and design*: AK Peters/CRC Press.
- OPAIN-S.A. (2019). *Plan Operativo del Aeropuerto Internacional El Dorado* (PLN-OPS-001). Retrieved from <https://www.opain.co/archivos/PLAN%20OPERATIVO%202019.pdf>
- OPAIN-S.A. (2020). Estadísticas del Aeropuerto. Retrieved from <https://eldorado.aero/comunicados/estadisticas-del-aeropuerto>
- Proceedings, O. O. A. E. A. (2018). *Air Travel Consumer Report*. Retrieved from <http://www.transportation.gov/airconsumer>
- Qlik. (2019). Data Analytics for Modern Business Intelligence. Retrieved from <https://www.qlik.com/us>
- Rollins, J. (2015). Why we need a methodology for data science. Retrieved from <https://www.ibmbigdatahub.com/blog/why-we-need-methodology-data-science>
- Sapp C., B. D., Antelmi J., Cook H., Craig T., Barot S. (2018). 2019 Planning Guide for Data and Analytics. Retrieved Oct, 5, 2018.
- SeatGuru. (2019). Fleet Information and Seat Maps. Retrieved from https://www.seatguru.com/airlines/United_Airlines/fleetinfo.php
- SGR-Comunicaciones. (2018). Aeroflot: primera aerolínea rusa en realizar seguimiento de equipaje. Retrieved from <http://www.aviacionnews.com/blog/2018/01/aeroflot-primer-aerolinea-rusa-en-realizar-seguimiento-de-equipaje/>
- Shneiderman, B. (2000). The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. *IEEE Symposium on Visual Languages, Proceedings*.
- SITA. (2018). *The Baggage Report 2018*. Retrieved from <https://www.sita.aero/resources/type/surveys-reports/baggage-report-2018>
- SITA. (2019a). Acerca de SITA. Retrieved from <https://es.sita.aero/about-us/who-we-are>
- SITA. (2019b). Baggage management. Retrieved from <https://www.sita.aero/solutions-and-services/solutions/baggage-management>
- SITA. (2019c). WorldTracer. Retrieved from <https://www.sita.aero/solutions-and-services/products/worldtracer>
- SKYbrary. (2020). ICAO Aerodrome Reference Code. Retrieved from https://www.aviationsafetywiki.org/index.php/ICAO_Aerodrome_Reference_Code
- Tableau. (2019). Business Intelligence and Analytics Software. Retrieved from <https://www.tableau.com/>

- van Zundert, H. (2010). *Off-Airport Baggage Check-in*. TU Delft. Master Thesis,
- Wikipedia. (2018). Sistema automático de tratamiento de equipajes. Retrieved from
https://es.wikipedia.org/wiki/Sistema_autom%C3%A1tico_de_tratamiento_de_equipajes
- Wikipedia. (2019a). METAR. Retrieved from <https://en.wikipedia.org/wiki/METAR>
- Wikipedia. (2019b). WorldTracer. Retrieved from
<https://en.wikipedia.org/wiki/WorldTracer>
- Wikipedia. (2020). International Civil Aviation Organization. Retrieved from
https://es.wikipedia.org/wiki/Organizaci%C3%B3n_de_Aviaci%C3%B3n_Civil_Internacional
- Zeinaly, Y., De Schutter, B., & Hellendoorn, H. (2013). *A model predictive approach for baggage handling systems*. Paper presented at the 16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013).