Technology

Thank you to our Sponsor

aMADEUS
Opening Remarks

Juan Ivan Martin
Head of Innovation
IATA
Airlines in the Data Science Era:
Borrowing Disruption From Other Industries

Alex Cosmas
Expert Associate Partner
McKinsey & Company
Airlines in the Data Science Era
Borrowing disruption from other industries

Alex Cosmas
Data is being produced at record volumes. However, blind reliance on mining increasing volumes of data is leading to (useless) information overload

“Petabytes allow us to say: ‘Correlation is enough’.”

- Chris Anderson, Editor-in-Chief
  Wired Magazine, June 23, 2008

The big data community is growing overly confident in data’s ability to explain everything

Real Fact #10: Mosquitoes are attracted to people who just ate bananas
Joseph
- Lives in Dallas
- Appears to fly American only to cities not served by Southwest
- Target Alice with discounts or bonus offers on flights to Southwest cities

Alice
- Lives in Singapore
- Purchases WiFi whenever it is available/rarely books flights without WiFi
- Joseph is searching for a ticket on SQ’s website. No WiFi-enabled flights are available. Incentivize Joseph to travel with bonus miles so that he does not book with a competing carrier

Charlie
- Lives in Dubai
- Needs to fly from DXB to BOM
- Always purchases business class fares for business travel
- Charlie is searching for an itinerary where business class is sold out. Incentivize Charlie to travel with an upgrade certificate to be used on future personal travel
Pandora Radio

explaining (and articulating) what humans cannot
Spotify’s *Taste Rewind*

extrapolate unobserved history (and present emotion) based on current behavior
Bayesian Diagnostics replicate the reasoning of seasoned technicians for live diagnosis.
Google DeepMind

outplay the human brain in multi-step action states, such as game-theoretic scenarios
U.S. Special Forces
downselect candidates before they fail, and prescribe the regimen so they don’t
TrueFit

eliminate friction of online apparel shopping using apparel DNA and consumer choice
Nutonian’s Eureqa

empirically learning the laws of nature, without (biased) theory as a starting point
Sports Analytics

Quantify everything and figure out the application later.
Fare differential

Incentivize booking of higher fare on similar quality itinerary

Itinerary Quality

Incentivize booking of less desirable routing

Market/Freq. Share

Incentivize booking of less popular carrier

Product Quality

Incentivize booking away from competitor’s higher quality product

Valuation of passengers’ disutility is dependent on each individual’s cognitive model
The data required for dozens of use cases are already captured along the various components of a passenger’s journey:

- **Ticket Purchase**
  - PNR Data: Origin, dest., routing, etc.
- **Seat Selection**
  - Premium seat purchase
- **Luggage**
  - Number of bags
  - Weight
- **Boarding**
  - Priority purchase
  - Boarding time
- **Ground Transportation**
  - Mode
  - Destination

**Search**
- Kind of tickets searched for (refundable, one-way)
- Available fares at time of purchase Ticket info viewed

**Changes**
- Change dates
- Change types (class, routing, etc.)

**Check-in**
- Check-in platform
- Check-in time

**Airport Activity**
- Travel path
- Purchases

**Onboard Ancillary**
- Wi Fi
- Food/beverage

**Reviews**
- Official complaints
- Social media/blog reviews
The Path to a Data Driven Airline

Ramki Ramaswamy
VP IT, Technology Integrations
JetBlue
Directions

- Identify the opportunities
- Identify the data sources
- Capture the relevant data and classify them
- Store and catalog the data
- Enable self-serve real time analytics
- Finish off with a feedback loop to the workflow
IDENTIFY THE OPPORTUNITIES

Value Driven Workflow Identification

- Commercial
- Operations
- Aircraft maintenance
- Crew management
IDENTIFY THE DATA SOURCES

- Multiple sources of data
  - Revenue systems
  - Crew/employee systems
  - Flight management systems
  - Flight planning systems
  - Maintenance systems
  - Third party ancillary sales

- Multiple types of data
  - Time series data
  - Pricing data
  - Availability data
  - Schedule data
  - Unstructured data interaction and news data
  - Weather, fuel pricing, etc.
CAPTURE THE RELEVANT DATA AND CLASSIFY THEM

- Transactional data
- Loyalty data
- PCI data
- Inventory data
- Unstructured data
STORE AND CATALOG THE DATA

- Size of the data
- Single source of truth
- Data and their relationships
- Loosely couple data
- Real time content
- Event and notification data
- Compliance and security
- Scalability
- Availability
- Accuracy & data quality
- History of data
FINISH WITH A FEEDBACK LOOP TO THE DECISION PROCESS

- Enriched data should feed the decision process in real time
- Generate new events
- Build out systems with loosely chained events
THANK YOU!
Airport which Puts Data Use at its Core

- Maurice Jenkins
  CIO, Miami International Airport
- Eric Knowles
  Senior Systems Analyst/Programmer, Miami International Airport
- Carlos Garcia
  Senior Computer Services Manager, Miami International Airport
PUTTING DATA AT THE CORE AT MIA

November 15th, 2017
“I’m focused on these statistics for good reason; at the end of the day, they represent the easiest and most reliable way to gauge the success of MIA and our general aviation airports.”

Emilio T. González
MIA Airport Director

“A tool that will significantly enhance our ability to view, assess and understand our operations.”

Joseph Napoli
MIA Chief of Staff & Senior Policy Advisor
Keys to Business Intelligence (BI) Project Success

- Support from administrators and project sponsor
- Access to complete and accurate data
- Meeting and buy-in with each business unit
  - Administration, Customer Service, Finance, Marketing
  - Operations, Properties & Concessions, Security, Maintenance & Construction
- Understand the needs of all stakeholders
- Identify Key Performance Indicators (KPIs)
- Focus on enhancing planning and efficiency
- Make the visualizations easy to understand and parameter based
- Core Operations targeted first, followed by low hanging fruit
Challenges

- Dealing with incomplete or bad data
- Integrating disparate data sources
- Gaining access to proprietary or sensitive data sources
- Changing the business processes to support the initiative
- Developing a culture of decision-making based on the data
- Choosing the right Business Intelligence (BI) tool
What is Tableau?

Tableau is an innovative approach to business intelligent applications. It is a complete data analytics platform that helps create visualizations using graphs, tables, and maps. It allows businesses to understand data and make informed decisions. It connects easily to various data sources, such as Oracle, SQL, Excel, and text files. Its efficiencies allow for "Real-Time" dashboard interaction.
Why Tableau at MIA

• It is a tool that can be used by non-technical professionals.
• It is not only easy to use, it is also interactive, real-time, parameter driven and effective enough to fulfill your highly professional needs.
• It is a user-friendly tool. Provides an easy-to-use drag and drop interface.
• Designed to be faster than other BI tools.
• Makes data accessible to anyone.
• Provides data drill down.
• Easily extracted to Excel for use in further analysis and presentations.
Our Achievements

We have built approximately 40 Dashboards for 8 Business Units.

Project and related training focused on Divisional managers and above.

We have been able to get fast answers to critical business questions.

Users can quickly analyze, visualize, and share information.

We have reduced the amount of time and effort compiling data from different sources.

We have helped our business executives save time and effort collecting and preparing information to make informed decisions.

Users can extract faster crucial facts from a large amount of data.
## BI Reports at MIA

**Economic Powerhouse · Industry Leader · International Brand**

<table>
<thead>
<tr>
<th><strong>• Administration</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>• Overtime</td>
</tr>
<tr>
<td>• HEAT Ticket Tracking</td>
</tr>
<tr>
<td>• Evaluations Tracking (TBD)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>• Customer Service</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>• Protocol Tracking</td>
</tr>
<tr>
<td>• Customer Service (TBD)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>• Finance</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>• Authorized Travel Tracking</td>
</tr>
<tr>
<td>• Aging (Receivable) *</td>
</tr>
<tr>
<td>• Aging (Payables)</td>
</tr>
<tr>
<td>• Revenue Budget</td>
</tr>
<tr>
<td>• Revenue Cash Basis Tracking</td>
</tr>
<tr>
<td>• Operating Budget *</td>
</tr>
<tr>
<td>• Reserved Maintenance Usage</td>
</tr>
<tr>
<td>• Reported Activity</td>
</tr>
<tr>
<td>• Project Controls Dashboard (TBD)</td>
</tr>
<tr>
<td>• Project Controls with Schedules (TBD)</td>
</tr>
<tr>
<td>• Grants (TBD)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>• Marketing</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>• Market Analysis *</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>• Operations</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>• Delayed Flights</td>
</tr>
<tr>
<td>• Flights per Hour by Alley</td>
</tr>
<tr>
<td>• Flights by Hour</td>
</tr>
<tr>
<td>• FIS Passengers by Hour</td>
</tr>
<tr>
<td>• Gate Utilization &amp; Details *</td>
</tr>
<tr>
<td>• International Baggage Efficiency</td>
</tr>
<tr>
<td>• ITBF Baggage Tracking</td>
</tr>
<tr>
<td>• Passengers per Hour by Concourse</td>
</tr>
<tr>
<td>• Passengers per Hour by Terminal</td>
</tr>
<tr>
<td>• Spot Utilization</td>
</tr>
<tr>
<td>• Countries Served *</td>
</tr>
<tr>
<td>• Daily Statistics</td>
</tr>
<tr>
<td>• AOA Decals</td>
</tr>
<tr>
<td>• Traffic Forecast/Comparison</td>
</tr>
<tr>
<td>• APC Kiosks</td>
</tr>
<tr>
<td>• CUSS &amp; Cute Usage</td>
</tr>
<tr>
<td>• Aircraft Movements</td>
</tr>
</tbody>
</table>

* Included in Demonstration
• Properties & Concessions
  • Agreements
  • Vendor Performance (TBD)
  • Vendor Compliance (TBD)
  • Concessions Inventory (TBD)

• Security
  • Badge Expiration
  • AOA Gates Activity (TBD)
  • Badges Issued / Fingerprints (TBD)

• Maintenance and Construction (TBD)
  • EAMS WO Tracking
  • Open Work Orders (EAMS)
  • MC7040 Project Tracking
  • Vehicle Fleet Management
  • ISO Certifications
  • Plan Review
  • WO/Req/PO Timeline

• Miscellaneous (TBD):
  • Accidents
  • Recruitments
  • Vacancies
  • Contracts
  • Personnel Retirement Report
  • Incidents
  • Public Records Request
  • Recycling

BI Reports at MIA
Economic Powerhouse • Industry Leader • International Brand
Operational BI Reports at MIA

Economic Powerhouse • Industry Leader • International Brand
ECONOMIC POWERHOUSE • INDUSTRY LEADER • INTERNATIONAL BRAND

THANK YOU!
<table>
<thead>
<tr>
<th>Country</th>
<th>Airport</th>
<th>Airline</th>
<th>Flights</th>
<th>Passengers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anguilla</td>
<td>Anguilla</td>
<td>General Aviation</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Argentina</td>
<td>Buenos Aires</td>
<td>American</td>
<td>374</td>
<td>81,414</td>
</tr>
<tr>
<td>Argentina</td>
<td></td>
<td>Argentinian</td>
<td>323</td>
<td>81,390</td>
</tr>
<tr>
<td>Argentina</td>
<td></td>
<td>Atlas Air</td>
<td>51</td>
<td>0</td>
</tr>
<tr>
<td>Argentina</td>
<td></td>
<td>Florida West</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Argentina</td>
<td></td>
<td>General Aviation</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Argentina</td>
<td></td>
<td>LAN</td>
<td>179</td>
<td>37,625</td>
</tr>
<tr>
<td>Argentina</td>
<td></td>
<td>LAN Cargo</td>
<td>26</td>
<td>0</td>
</tr>
<tr>
<td>Argentina</td>
<td></td>
<td>Material</td>
<td>55</td>
<td>0</td>
</tr>
<tr>
<td>Argentina</td>
<td></td>
<td>UPS</td>
<td>57</td>
<td>0</td>
</tr>
<tr>
<td>Arabia</td>
<td>Arabian</td>
<td>American</td>
<td>517</td>
<td>66,455</td>
</tr>
<tr>
<td>Arabia</td>
<td>Arabian</td>
<td>Jazeera</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>Arabia</td>
<td>Arabian</td>
<td>Iran Air</td>
<td>117</td>
<td>17,264</td>
</tr>
<tr>
<td>Arabia</td>
<td>Arabian</td>
<td>Saudia</td>
<td>77</td>
<td>6,622</td>
</tr>
<tr>
<td>Arabia</td>
<td>Arabian</td>
<td>United</td>
<td>2</td>
<td>318</td>
</tr>
<tr>
<td>Arabia</td>
<td>Arabian</td>
<td>World Atlantic</td>
<td>1</td>
<td>80</td>
</tr>
<tr>
<td>Australia</td>
<td>Yeeld, Australia</td>
<td>General Aviation</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Australia</td>
<td>Yeeld, Australia</td>
<td>IBC Airways</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Austria</td>
<td>Vienna</td>
<td>General Aviation</td>
<td>1</td>
<td>15</td>
</tr>
<tr>
<td>Bahamas</td>
<td>Andros Town</td>
<td>General Aviation</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Bahamas</td>
<td>Eleuthera</td>
<td>General Aviation</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Bahamas</td>
<td>Eleuthera</td>
<td>General Aviation</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Bahamas</td>
<td>Eleuthera</td>
<td>Eagle</td>
<td>163</td>
<td>7,007</td>
</tr>
<tr>
<td>Bahamas</td>
<td>Eleuthera</td>
<td>General Aviation</td>
<td>16</td>
<td>28</td>
</tr>
<tr>
<td>Bahamas</td>
<td>Eleuthera</td>
<td>IBC Airways</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Bahamas</td>
<td>Eleuthera</td>
<td>Eagle</td>
<td>628</td>
<td>25,124</td>
</tr>
<tr>
<td>Bahamas</td>
<td>Eleuthera</td>
<td>General Aviation</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Bahamas</td>
<td>Eleuthera</td>
<td>IBC Airways</td>
<td>88</td>
<td>195</td>
</tr>
<tr>
<td>Bahamas</td>
<td>Eleuthera</td>
<td>MMH&amp;cargo</td>
<td>117</td>
<td>0</td>
</tr>
<tr>
<td>Bahamas</td>
<td>Eleuthera</td>
<td>Skyway</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Bahamas</td>
<td>Eleuthera</td>
<td>Swift Air</td>
<td>26</td>
<td>908</td>
</tr>
<tr>
<td>Bahamas</td>
<td>Eleuthera</td>
<td>Bahamas Airlines</td>
<td>3</td>
<td>262</td>
</tr>
<tr>
<td>Bahamas</td>
<td>Eleuthera</td>
<td>Eagle</td>
<td>277</td>
<td>13,390</td>
</tr>
<tr>
<td>Bahamas</td>
<td>Eleuthera</td>
<td>General Aviation</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Bahamas</td>
<td>Eleuthera</td>
<td>World Atlantic</td>
<td>2</td>
<td>118</td>
</tr>
</tbody>
</table>
What Can You do with BI – Gate Utilization
What Can You do with BI - Budget

Budget Dashboard
Report Date: 9/13/2017 10:51:45 AM

Percent of Budget Used vs Percent of Fiscal Year Elapsed

Budget Summary by Organization

Adjusted Budget & Remaining Budget

Organization Group
Division
Account Type
- (All)
- Capital Expenditures
- Changes in County Surplus
- Cost of Goods Sold
- Debt Service
- Fees
- General & Admin Expenses
- Insurance
- Outside Contracts
- Overtime
- Registration
- Salaries
- Temporary Services
- Transfers to Other Funds
- Travel
- Utilities

Cancel   Apply
What Can You do with BI – AR Aging
What Can You do with BI – Marketing Analysis

Market Analysis
Date Range: 11/3/2017 5:55:37 PM
* Passenger counts received from the airlines on the day of operation.

Traffic Breakdown

Region | Country | Airport Name | Airline Name | Flights | Passengers
--- | --- | --- | --- | --- | ---
Africa | Cape Verde | Cape Verde | General Aviation | 1 | 0
| | | | General Aviation | 1 | 2
| | | | Western Global Aviation | 2 | 0
| | Total | | | 4 | 2
| Central African Republic | Brla, CF | General Aviation | 1 | 0
| | Total | | | 1 | 0
| Gabon | Libreville, Gabon | General Aviation | 1 | 0
| | Total | | | 1 | 0
| Morocco | Marrakech | KS Aviation | 2 | 36
| | Total | | | 2 | 36
| Senegal | Dakar | Atlas Air | 2 | 0
| | Total | | | 2 | 0
| Total | | | | 10 | 38
Caribbean | Anguilla | Anguilla | General Aviation | 11 | 12
| | | | General Aviation | 14 | 18
| | Total | | | 25 | 30
Antigua and Barbuda | Antigua | General Aviation | 15 | 26
| | | | Air Canada | 1 | 0
| | | | American | 728 | 93,335
| | | | AmerisJet | 47 | 0
| | | | Delta Air Lines | 1 | 0
| | | | General Aviation | 11 | 43
| | | | Miami Air | 2 | 0
| | | | Northern Air Cargo | 2 | 0
| | | | World Atlantic | 4 | 250

Flights

Passengers
The Future is Now: All About Big Data

Natascha Maegdefessel
Lead Data Scientist
ZeroG
THE FUTURE IS NOW
All about Big Data
the future is NOW? what does that mean?

O'HARA HEADSET

lol, youre right, your 'future' is our present, mom!

HOLODECK

BEAM US UP, SCOTTY
here's our view on the future!
what does that have to do with Big Data?? 😐😐

For all of this, you have to teach a machine to understand, to recognize and to feel you
For example, Big Data learning techniques teaches computers to recognize all this things immediately and in once?

Again: what does that have to do with Big Data???? 😐
A machine learns from data! BUT, even for a machine that is a messy problem in the first place.

LEARNING FROM BIG DATA STARTS WITH A MESSY PROBLEM
Also a very smart machine has started like a newborn. It is only by far faster and is able to take more in account at once.

Baby
That is a face

Teenager
That is a female face

Adult
That is Natascha’s face

Genius
That is the face of Natascha’s twin
#zeroG DeepSky

- Uses modern AI learning techniques (Re-Enforcement Learning)
- To teach the computer to play successfully the airline simulation game “Air Bucks”
- Training on a cluster within the Microsoft Azure Cloud
- Duration 6 month
DeepSky

- for creating e.g. an optimal schedule, concerning
- Passenger acceptance
- Fleet, Tail and Crew Assignment
- Delay Robustness

...and all that in once

Currently all this is done separately and in many cases in different departments.
Back in the days...

Airlines used to be geniuses concerning Big Data and Data Science.  
- They have been pioneers concerning digitalization of data and processes  
- Some data scientific methods has been developed especially for them  
- They established business processes performing great in the past and still working today - somehow

But well-established business processes are hard to transform into modern Processes fulfilling the requirements of our digital 21st century.

 Today airlines has fallen back in between the teenager and adult stage concerning Big Data and Data Science.
Live long…

 Drops

…and prosper!

…whatever that means.
Networking Coffee Break
Thank you to our Sponsor

INFARE
Leveraging External Data: The Value of Data Integration

Willie Cecil
Director, Aviation Uptake
How to Derive Value From Big Data: A Practical Case

Roy Hasson

Business Development, Analytics
Amazon Web Services
The Value of Data

How to get insights from big data: Qantas Case

Roy Hasson, Business Development Manager, Analytics, AWS

Miami, 16th November 2017
We combine data to make it actionable....We’re doing that together with Amazon, because there is only one company that we can do this with which gives us the reliability, scale, and performance we need.

Jeroen Tas
CEO, Healthcare Informatics Solutions and Services
Data is of no use if you can’t get actionable insights.
Enters **AWS**: Cloud Built for Big Data

- Agility
- Scalability
- Broadest and Deepest Capabilities
- Low Cost
- Get to Insights Faster
- Data Migrations Made Easy
At AWS, we know how to collect, store and analyze data
At AWS, we know how to get **insights** from data
How to build a **Travel Data Platform**

1. **Ingest data from multiple sources**
2. **Data storage without silos enable greater collaboration**
3. **Facilitate data exploration and machine learning**
4. **Integrate with existing business intelligence tools**
AWS Professional Services worked with Qantas to design, optimize and operationalize a high volume and critical real-time streaming platform. The Customer Agility Platform (CAP) is a new technology capability to provide (timely) information and generate insights to support the Qantas business to better engage with our customers, therefore improve customer experience & profitability. Combination of AWS and open source made this system practical, previous evaluations using on premise and proprietary software proved cost prohibitive and hard to scale.

David Hockridge,
Manager, Business Insight

Australia’s largest airline with annual turnover of over AS$16.2bn. Carried over 51 million passenger journeys in 2016

Unified customer ID and Group Segmentation tagging across Qantas Group: QFF, Non-QFF & Jetstar

- Real time access to customer bookings and journey information
- Booking change event notification engine

1.5M msg/day

Average: 20 msg/sec  Peak**: 500 msg/sec

Average: 80 KB per msg Max: 1.2 MB per msg

Highly resilient and reliable (Zero message loss)

* Compared to on-premises
** During recovery scenarios
Qantas: Business Opportunities

Near Real Time Data Updates
- Flights
- Bookings
- Disruptions
- Ancillary services
- Operational (Boarding / Check-in)

Customer profiles with history, metrics and preferences

Higher Customer Engagement with potential for:
- Real time marketing
- Personalize Sales
- Real Time notification
- Flight delays detection
- Re-accommodation
- Operation optimization
Conclusions

• The right tools get the best insights
• AWS can help driving business outcome
• Work the Amazon Way

innovate at startup speed
reduce risk
focus on differentiation

https://aws.amazon.com/big-data/
Thank you!
Come and find us if you want to know more…
A Case Study From a Pioneer in Predictive Analytics

Josef Habdank
Chief Data Scientist
Infare
Advanced prediction methods on hundreds of billions of Airfares using Apache Spark framework

IATA Aviation Data Symposium
Miami, USA 15-16 November 2017

Josef Habdank
Chief Data Scientist & Data Platform Architect

jha@infare.com
https://www.linkedin.com/in/jahabdank
Leading provider of Airfare Intelligence Solutions to the Aviation Industry

245 Airlines & 20 Airports, using our system daily

Collect 2 billion distinct airfares daily

~1 trillion airfares in the database
Prediction with Machine Learning | price evolution in Pharos

Forecast

D83091, CPH > LON
Prediction with Machine Learning | minimum future price in Pharos

7 days Forecast
How can this be done

- BigData framework: Apache Spark
- Online learning: training machine learning models real time
- Particle Predictor (Markov Chain Monte Carlo):
  - Price Prediction
  - Anomaly prediction
How can this be done

Apache Spark is an in-memory processing framework very popular for Machine Learning on BigData. It is being used by giants of the BigData industry such as Google, Facebook, IBM and many others.

Google Trends for “BigData” and “Apache Spark”

Google Trends for “PySpark” and “Python Spark”
Nature of AirFare prediction

- Flight prices are set using revenue management systems which internally use a finite state space.
- Price prediction can be thought as a state classification problem in the future time.
- At each time point there is a hidden posterior probability that needs to be estimated (probability of price remaining the same state, or go to other states).
Prediction using Particle Predictor (Markov Chain Monte Carlo)

- State space systems are well described by Markov Chains
- It estimates hidden non-normal posterior probability (the probability distribution of the next step) by creating histograms based on historical data for each state
- Prediction done by repetitively doing simulation of the Markov Chain (in our case 100 simulations give very satisfactory results)
- Produces a probability cloud that can be used to get highest likelihood path, as well as confidence intervals and anomalies
Particle Predictor in practice

- Particle predictor capable of predicting world wide prices can easily have tens of millions of PDFs which need to be stored in a massive DataBase.

- Each time one step prediction is done, the system needs to lookup the estimated posterior probability from the DataBase.

- Predicting price 15 days ahead, by doing 100 simulations results in 1500 database lookups among tens of millions of stored PDFs.

- Sub second time for prediction can be done but is very hard, need massive in memory database.
Particle Predictor in practice

good accuracy (28 day prediction)

anomaly predicted, manual price change
Future research in Machine Learning | real time estimation of world wide demand

Airline’s own historical and current demand curves

Historical and current market prices

Current and future demand curves

We are looking for partner Airlines to pilot the research project. Get in touch to know more!
Thank you!

Josef Habdank
Chief Data Scientist & Data Platform Architect

jha@infare.com
https://www.linkedin.com/in/jahabdank
Using Statistics to get From Data to Knowledge

Jesper Venema
Senior Statistician
IATA
IATA Statistics

Jesper Venema
Senior Statistician, IATA
IATA Statistics Program

• Supported by app. 260 airlines
  • All airline types
  • Global coverage (app. 95%)
  • Representation of all regions and major economies
  • Also non-members
IATA Air Transport Statistics

• Most timely real-sector statistics with global coverage
• High industry coverage (app. 95%)
• Referred globally as a predictor for economic down- or up-turns, international commerce
Passenger Traffic Growing Steadily

Industry RPKs (billion per month)

Sources: IATA Economics, IATA Monthly Statistics
Air Cargo at a High

Industry FTKs (billion per month)

Sources: IATA Economics, IATA Monthly Statistics
Passenger Yields

Global average yield, US$ constant exchange rate (Jan 2011)

Global average

US$ per RPK, seasonally adjusted
Cargo Yields improve but profitability a challenge

Source: IATA Economics, IATA CASS, Platts
Industry-wide, increasing load factors

Passenger load factor, seasonally adjusted
Freight load factor, seasonally adjusted

% of ASKs
% of AFTKs
AIR PASSENGER MARKET ANALYSIS

May 2016

Robust growth in 2016 so far, but ongoing signs of moderation

- The global passenger market has made a robust start to 2016, growing in line with its 10-year average...
- But there are ongoing signs that passenger growth is slowing down a year...
- The domestic trade market remains the one performer, with volumes up more than 3% in the first half of 2016.
- The industry appears to have stabilised the recent easing in the (seasonally adjusted) trend decline.

A robust start to 2016 as a whole...

The global air passenger market has made a robust start to 2016 so far. Nonetheless, the 10-year trend remains an essential indicator for passenger growth (TFAs) in the first half of 2016 has been followed by the same trend. But even after surpassing the same data in February, we estimate that traffic still grew by 3.3% this data is too low. This trend has been supported by the sharp rise in the past decade...

...but growth moderation has occurred...

Nonetheless, following a bumper year in 2015, there are ongoing signs that passenger growth has started to slow. Annual growth in TFAs continued weakly at 4.0% in May — its slowest pace since January 2016.

Chart 1 — Air passenger volume growth and global business confidence

Air passenger market overview — May 2016

Aircraft market overview — May 2016

AIR FREIGHT MARKET ANALYSIS

May 2016

Upward momentum in FTAs has stalled in 2016

- Annual growth in air freight traffic fell to 3.5% in March 2016.
- The upward momentum that was established in 2015 has dropped in the first quarter of this year. (See Chart 1.)

Volume growth rates are an important indicator for freight volumes, which dropped in the US market this year. In April, freight volumes fell 0.3% compared to March 2015. (See Chart 1.)

While freight rates remain strong, the overall impact on the air cargo market is positive. The headline figures for the first quarter of the year show that freight volumes fell 0.3% compared to the same period last year. (See Chart 2.)

Hotels and other travel-related businesses have broadly trended higher since the start of 2016, and the relationship between global tourism and hotel beds remains in line with 2015.

Chart 2 — Air freight growth vs. hotel market

Air Freight Market Analysis — May 2016

Aircraft Market Analysis — May 2016

*1% of change in FTKs 2016 vs. 2015. *2% of change in hotel market

Air Passenger Market Analysis — May 2016

*1% of change in FTKs 2016 vs. 2015. *2% of change in hotel market
IATA Statistics

IATA at the forefront

• Data sources and repositories uniquely available to IATA
• Direct airline reporting
  • Segment origin-destination traffic data, passengers and cargo
  • Ticketing data
  • Financial data
  • Employment, Fleet
• Available Settlement data (ticketing, cargo)
• Industry knowledge and expertise within IATA
• Information technology
IATA Statistics

Making industry statistics is not just adding data

• application of statistical definitions and conventions
• data collection mechanisms
• validation
• data dissemination solutions
• statistical techniques
• IT solutions
• industry knowledge and expertise
IATA Statistics

Making industry statistics is not just adding data

• Just data do not provide the entire story, many traps and holes, quantity is not quality
• Most available data stem from administrative records
• Coverages, concepts and definitions, people change, systems change
• When combining multiple data repositories, these issues grow exponentially
New techniques and statistical approaches, combining unique data repositories with industry expertise

Examples:
• Validation by outlier detection
• Market size estimations
• Forecasting
• New statistics: passenger traffic by nationality
Validation by Outlier Detection

• On-Flight Origin-Destination data
  • Used for variety of products
  • App. 38,000 records per month
  • Validation can only be done contextually
    - Intra-source
      • cross airline
      • historical
    - Inter-source
      • schedules databases
  • Mistakes usually involve an entire filing
Validation by Outlier Detection

- Implied load-factors
  - # of incidents high / low
- Growth rates assessment
  - Comparison with historical behavior
Validation by Outlier Detection

Inter-source growth rates assessment – outliers above accepted standard error interval

Acceptance range based on 4 million observations, 15 years of historical data
Validation by Outlier Detection

• Implied load-factors
  • # of incidents high / low

• Growth rates assessment
  • Comparison with historical cross-airline patterns
  • Comparison with schedules
    • # of incidents outside of range
Validation by Outlier Detection

Intra-source growth rates assessment – outliers above accepted standard error interval

Acceptance range based on 4 million observations, 15 years of historical data
Validation by Outlier Detection

Intra-source growth rates assessment – outliers above accepted standard error interval

Acceptance range based on 4 million observations, 15 years of historical data
Validation by Outlier Detection

• Implied load-factors
  • # of incidents high / low

• Growth rates assessment
  • Comparison with historical cross-airline patterns
  • Comparison with schedules
    • # of incidents outside of range
    • Weight score of these incidents
# Validation by Outlier Detection

## Score card concept

<table>
<thead>
<tr>
<th>Type</th>
<th>Airline.Code</th>
<th>201709</th>
<th>201708</th>
<th>201707</th>
<th>201706</th>
</tr>
</thead>
<tbody>
<tr>
<td>OFODS Submission</td>
<td>AYG</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>repeating_FRT</td>
<td>CHO</td>
<td>Repeated</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LoadFactorHigh</td>
<td>MKS</td>
<td>24</td>
<td>22</td>
<td>26</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>RJM</td>
<td>15</td>
<td>8</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>PKL</td>
<td>54</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>LoadFactorHigh</td>
<td>CMN</td>
<td>24</td>
<td>18</td>
<td>11</td>
<td>18</td>
</tr>
<tr>
<td>LoadFactorLow</td>
<td>PKL</td>
<td>39</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>LoadFactorLow</td>
<td>KDI</td>
<td>10</td>
<td>10</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Outlier</td>
<td>TVX</td>
<td>68</td>
<td>67</td>
<td>71</td>
<td>63</td>
</tr>
<tr>
<td>Outlier</td>
<td>PKL</td>
<td>28</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Outlier</td>
<td>CWS</td>
<td>20</td>
<td>16</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>Outlier</td>
<td>IWK</td>
<td>10</td>
<td>2</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Weight</td>
<td>TVX</td>
<td>1.6</td>
<td>1.5</td>
<td>1.6</td>
<td>1.7</td>
</tr>
<tr>
<td>Weight</td>
<td>PKL</td>
<td>3.3</td>
<td>1</td>
<td>1.1</td>
<td>1.2</td>
</tr>
</tbody>
</table>
## Validation by Outlier Detection

### Score card concept

<table>
<thead>
<tr>
<th>Type</th>
<th>Airline.Code</th>
<th>201709</th>
<th>201708</th>
<th>201707</th>
<th>201706</th>
</tr>
</thead>
<tbody>
<tr>
<td>OFODS Submission</td>
<td>AYG</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>repeating_FRT</td>
<td>CHO</td>
<td>Repeated</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LoadFactorHigh</td>
<td>MKS</td>
<td>24</td>
<td>22</td>
<td>26</td>
<td>28</td>
</tr>
<tr>
<td>LoadFactorHigh</td>
<td>RJM</td>
<td>15</td>
<td>8</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>LoadFactorHigh</td>
<td>PKL</td>
<td>54</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>LoadFactorHigh</td>
<td>CMN</td>
<td>24</td>
<td>18</td>
<td>11</td>
<td>18</td>
</tr>
<tr>
<td>LoadFactorLow</td>
<td>PKL</td>
<td>39</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>LoadFactorLow</td>
<td>KDI</td>
<td>10</td>
<td>10</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Outlier</td>
<td>TVX</td>
<td>68</td>
<td>67</td>
<td>71</td>
<td>63</td>
</tr>
<tr>
<td>Outlier</td>
<td>PKL</td>
<td>28</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Outlier</td>
<td>CWS</td>
<td>20</td>
<td>16</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>Outlier</td>
<td>IWK</td>
<td>10</td>
<td>2</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Weight</td>
<td>TVX</td>
<td>1.6</td>
<td>1.5</td>
<td>1.6</td>
<td>1.7</td>
</tr>
<tr>
<td>Weight</td>
<td>PKL</td>
<td>3.3</td>
<td>1</td>
<td>1.1</td>
<td>1.2</td>
</tr>
</tbody>
</table>
# Validation by Outlier Detection

## Score card concept

<table>
<thead>
<tr>
<th>Type</th>
<th>Airline.Code</th>
<th>201709</th>
<th>201708</th>
<th>201707</th>
<th>201706</th>
</tr>
</thead>
<tbody>
<tr>
<td>OFODS Submission</td>
<td>AYG</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>repeating_FRT</td>
<td>CHO</td>
<td>Repeated</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LoadFactorHigh</td>
<td>MKS</td>
<td>24</td>
<td>22</td>
<td>26</td>
<td>28</td>
</tr>
<tr>
<td>LoadFactorHigh</td>
<td>RJM</td>
<td>15</td>
<td>8</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>LoadFactorHigh</td>
<td>PKL</td>
<td>54</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>LoadFactorHigh</td>
<td>CMN</td>
<td>24</td>
<td>18</td>
<td>11</td>
<td>18</td>
</tr>
<tr>
<td>LoadFactorLow</td>
<td>PKL</td>
<td>39</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>LoadFactorLow</td>
<td>KDI</td>
<td>10</td>
<td>10</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Outlier</td>
<td>TVX</td>
<td>68</td>
<td>67</td>
<td>71</td>
<td>63</td>
</tr>
<tr>
<td>Outlier</td>
<td>PKL</td>
<td>28</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Outlier</td>
<td>CWS</td>
<td>20</td>
<td>16</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>Outlier</td>
<td>IWK</td>
<td>10</td>
<td>2</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Weight</td>
<td>TVX</td>
<td>1.6</td>
<td>1.5</td>
<td>1.6</td>
<td>1.7</td>
</tr>
<tr>
<td>Weight</td>
<td>PKL</td>
<td>3.3</td>
<td>1</td>
<td>1.1</td>
<td>1.2</td>
</tr>
</tbody>
</table>
## Validation by Outlier Detection

### Score card concept

<table>
<thead>
<tr>
<th>Type</th>
<th>Airline.Code</th>
<th>201709</th>
<th>201708</th>
<th>201707</th>
<th>201706</th>
</tr>
</thead>
<tbody>
<tr>
<td>OFODS Submission</td>
<td>AYG</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>repeating_FRT</td>
<td>CHO</td>
<td></td>
<td></td>
<td>Repeated</td>
<td></td>
</tr>
<tr>
<td>LoadFactorHigh</td>
<td>MKS</td>
<td>24</td>
<td>22</td>
<td>26</td>
<td>28</td>
</tr>
<tr>
<td>LoadFactorHigh</td>
<td>RJM</td>
<td>15</td>
<td>8</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>LoadFactorHigh</td>
<td>PKL</td>
<td>54</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>LoadFactorHigh</td>
<td>CMN</td>
<td>24</td>
<td>18</td>
<td>11</td>
<td>18</td>
</tr>
<tr>
<td>LoadFactorLow</td>
<td>PKL</td>
<td>39</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>LoadFactorLow</td>
<td>KDI</td>
<td>10</td>
<td>10</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Outlier</td>
<td>TVX</td>
<td>68</td>
<td>67</td>
<td>71</td>
<td>63</td>
</tr>
<tr>
<td>Outlier</td>
<td>PKL</td>
<td>28</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Outlier</td>
<td>CWS</td>
<td>20</td>
<td>16</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>Outlier</td>
<td>IWK</td>
<td>10</td>
<td>2</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Weight</td>
<td>TVX</td>
<td>1.6</td>
<td>1.5</td>
<td>1.6</td>
<td>1.7</td>
</tr>
<tr>
<td>Weight</td>
<td>PKL</td>
<td>3.3</td>
<td>1</td>
<td>1.1</td>
<td>1.2</td>
</tr>
</tbody>
</table>
Validation by Outlier Detection

• Quick identification of problem spots
• Flexible approach combining multiple validation criteria at an level of complexity
• Combination of statistical techniques and industry expertise
• Enhanced quality while time saving
Market Size Estimations

• No single data source or repository provides market totals
• Complementing data sources
  • Overlap
  • Each having different
    • Coverages
    • Biases
    • Concepts
Market Size Estimations

Cargo Market Totals:

• Traffic by flight segment (app 16,000 monthly records)
• Traffic by origin-destination data (app 1.5 million monthly records)

Application of proportions, while avoiding biases and application of statistical significance rules
Market Size Estimations

For a given airline:

\[
\text{Traffic A to P: segment [H-P] } \times \alpha
\]
Market Size Estimations

For a given airline:

\[ \alpha + \kappa \]

\[ \beta + \lambda \]

\[ \gamma + \mu \]
For a given airline:

Inferior statistical coverage in B: 
α becomes *preferred*;  
κ = segment [H-P] * (1-α)
Market Size Estimations

Multiple hubs:
Traffic A to P:
segment [H-P] * (α-δ) +
segment [G-P] * (δ)
Forecasting

• Ticketing data repositories contain a wealth of information for forward travel
• Short-term: up to three months
• Ability to detect trends and exceptions by using historical data
Forecasting

Booking Curve Example (based on 6,000 bookings on a particular day)
Forecasting

Booking Curve Example – Scenario of Extended Booking Trend

Historical Booking Curve  Actuals  Forecast
Forecasting

Booking Curve Example – Scenario of Prolonged Incident

Historical Booking Curve  Actuals
Forecasting

Booking Curve Example – Scenario of Prolonged Incident
Forecasting

Booking Curve Example – Scenario of Incident with Recovery

Historical Booking Curve
Actuals
Forecasting

Booking Curve Example – Scenario of Incident with Recovery
Forecasting

• Ticketing data repositories contain a wealth of information for forward travel
• Short-term: up to three months
• Ability to detect trends and exceptions by using historical data
Passenger Traffic by Nationality

• Product launched in partnership with DIIO and M1ndset
• Large ticketing databases available
  • Journey point of origin
  • Ticket point of sale
• Combination with additional data sources
  • Demographic repositories
  • Immigration data
  • Profiling using historical data sets
    • Expats and propensities to travel
    • Border airports
Passenger Traffic by Nationality

Border Airports
Passenger Traffic by Nationality

- Estimation by proportions while able to adjust for atypical patterns

**Point of Origin**
Ticketing Database for 1500 airports worldwide

- 3 million records per month

**Factorial adjustments:**
- Demographic and migration repositories
- Propensity to fly or “expat’ adjustments
- Airport specific adjustment
- Country specific adjustment
IATA Statistics

Conclusions

• IATA is able to leverage
  • access to unique data sources
  • industry knowledge and expertise assessing and validating available data, making use of complementing data sources
  • statistical techniques and IT sources

• IATA uniquely positioned to provide industry statistics
What is the Future of Security for Data

Kevin Iverson

CTO

Journera
What is the Future of Security for Data
Me

• Kevin Iverson
  kiverson@journera.com

• CTO @ Journera
  • Real-time data exchange backed by 6
    major airlines and hotels

• Background
  • 9 years @ Orbitz
  • 5 years @ various small-big companies
    focused on using data in Real Estate,
    Consumer Reviews, and Commercial
    Construction and A
Data is the new Currency

- Data as Currency is real – Top 5 Market Cap Companies
- Data is the tool to unlock experiences

But how we protect data is not working
It's time to change our approach

• Zero Trust
  • Stop assuming any part of your environment can be trusted

• Principle of Least Plain Text
  • Stop assuming that it's ok to keep everything in plaintext
Zero Trust

• Move beyond focus on perimeter security to protecting everything

• Ways to get there
  • Encryption in transit and at rest
  • Fine grained network access
  • Authenticate everything
  • Infrastructure as code with audits
  • Short lived services (Pets vs Cattle)
  • Audit logs
Principle of Least Plain Text (PoLPT)

• Keep in plain text only as a last result

• Some areas to consider:
  • Hash at field level as much as possible (hint: more than passwords)
  • Cryptography as a core environment service
  • Privacy preserving data analysis
  • Data retention evaluation
Questions?
Wrap-up and Closing

Thank you to our Sponsor

AMADEUS
Networking Dinner
Join us tonight!
19:45 – 23:00
In this room!