|  |
| --- |
| EPAM Systems, RD Dep. |
| Airline Company |

Image goes here.

Contents

[1. Business Description 3](#_Toc412572569)

[1.1. Business background 3](#_Toc412572570)

[1.2. Problems because of poor data management 3](#_Toc412572571)

[1.3. Benefits from implementing a Data Warehouse 3](#_Toc412572572)

[2. Dimensions of a Business 3](#_Toc412572573)

[3. Logical Scheme 3](#_Toc412572574)

[4. Data Flow 3](#_Toc412572575)

[5. Fact Table Partitioning Strategy 3](#_Toc412572576)

[6. Strategy of Parallel Load 3](#_Toc412572577)

[7. Report Layouts 3](#_Toc412572578)

# Business Description

## Business background

* + **Business background**

DWH will provide research department of the all airports and companies schedule with flight performance information (departures, arrivals, etc.).

* + **Problems because of poor data management**

Today it’s hard to know/analyze data about flight performance.

* + **Benefits from implementing a Data Warehouse**

As a result the data will have performance statistics that can be analyzed to improve companies and airports schedule and gain more passengers by static time flights.

# Dimensions of a Business

1. **Time Dim**   
   year, quarter, month, day of month, day of week, flight date(yyyymmdd)
2. **Airline Dim**  
   Unique Carrier Code, Airline ID, Carrier, Tail Number, Flight Number
3. **Origin Dim**  
   Origin Airport ID, Origin Airport Seq ID, Origin City Marke tID, Origin Airport, Origin City Name, Origin State Code, Origin State Name, Origin Airport World Area Code
4. **Destination Dim**

DestAirportID Destination Airport, Airport ID. An identification number assigned by US DOT to identify a unique airport. Use this field for airport analysis across a range of years because an airport can change its airport code and airport codes can be reused.

DestAirportSeqID Destination Airport, Airport Sequence ID. An identification number assigned by US DOT to identify a unique airport at a given point of time. Airport attributes, such as airport name or coordinates, may change over time.

DestCityMarketID Destination Airport, City Market ID. City Market ID is an identification number assigned by US DOT to identify a city market. Use this field to consolidate airports serving the same city market.

Dest Destination Airport

DestCityName Destination Airport, City Name

DestState Destination Airport, State Code

DestStateFips Destination Airport, State Fips

DestStateName Destination Airport, State Name

DestWac Destination Airport, World Area Code

1. **Departure Dim**

CRSDepTime (CRS Departure Time (local time: hhmm))

DepTime (Actual Departure Time (local time: hhmm))

DepDelay (Difference in minutes between scheduled and actual departure time. Early departures show negative numbers.)

DepDelayMinutes (Difference in minutes between scheduled and actual departure time. Early departures set to 0.)

DepDel15 (Departure Delay Indicator, 15 Minutes or More (1=Yes))

DepartureDelayGroups (Departure Delay intervals, every (15 minutes from <-15 to >180))

DepTimeBlk (CRS Departure Time Block, Hourly Intervals)

TaxiOut (Taxi Out Time, in Minutes)

WheelsOff (Wheels Off Time (local time: hhmm))

1. **Arrival Dim**

WheelsOn (Wheels On Time (local time: hhmm))

TaxiIn (Taxi In Time, in Minutes)

CRSArrTime (CRS Arrival Time (local time: hhmm))

ArrTime Actual Arrival Time (local time: hhmm)

ArrDelay (Difference in minutes between scheduled and actual arrival time. Early arrivals show negative numbers.)

ArrDelayMinutes (Difference in minutes between scheduled and actual arrival time. Early arrivals set to 0.)

ArrDel15 (Arrival Delay Indicator, 15 Minutes or More (1=Yes))

ArrivalDelayGroups (Arrival Delay intervals, every (15-minutes from <-15 to >180))

ArrTimeBlk (CRS Arrival Time Block, Hourly Intervals)

1. **Cancellations and Diversions**

Cancelled (Cancelled Flight Indicator (1=Yes)),

Cancellation Code (Specifies The Reason For Cancellation),

Diverted (Diverted Flight Indicator (1=Yes))

1. **Flight summaries** (or it can be as a part of fact table)

CRSElapsedTime CRS (Elapsed Time of Flight in Minutes),

ActualElapsedTime (Elapsed Time of Flight in Minutes),

Flight Time (in Minutes),

Number of Flights,

Distance,

Distance Group (Distance Intervals, every 250 Miles, for Flight Segment)

# Logical Scheme

# Data Flow

# Fact Table Partitioning Strategy

# Strategy of Parallel Load

# Report Layouts