



UNIVERSIDAD
DE MÁLAGA

Master en Advanced Analytics on Big Data

Knowledge Area: Big Data

Task 5: Deployment and use of an analytic system

Author: Juan Rafael Caro Romero

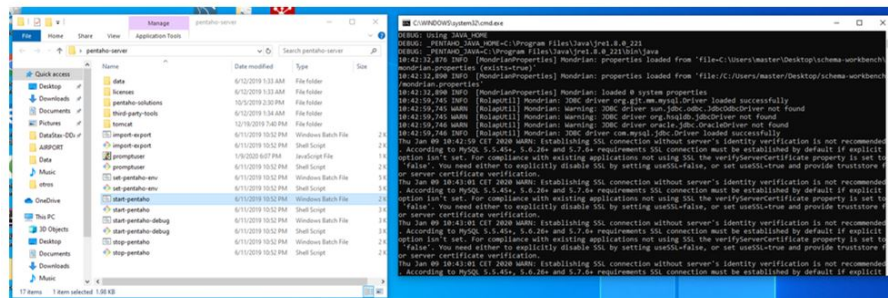
Juan Morales Conde

Martín Blazquez Moreno

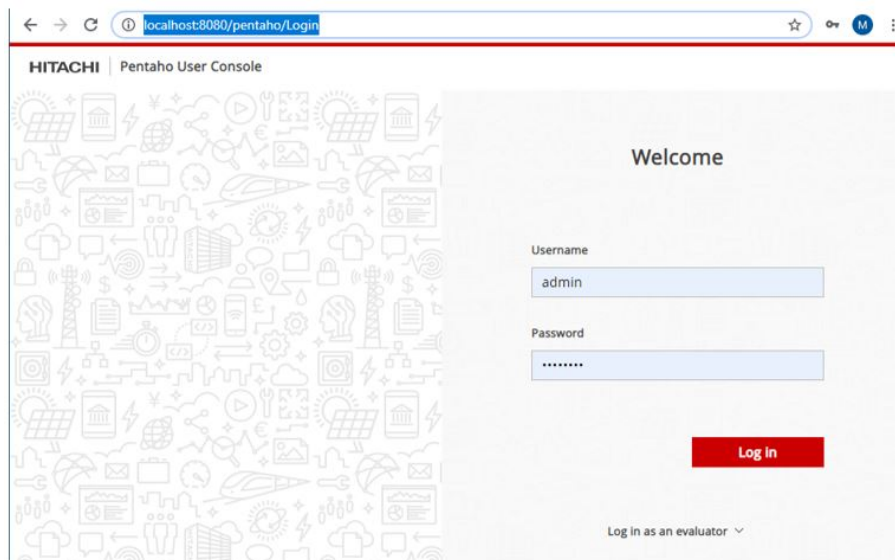
Module: Data Analytics

Málaga, 12th January 2020

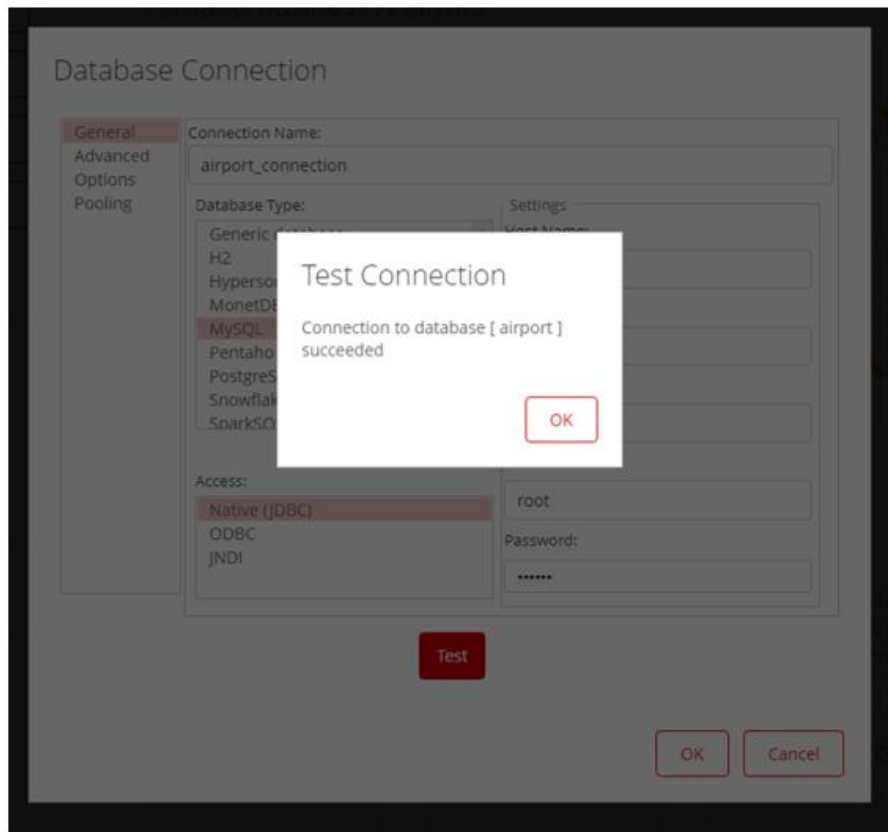
To start Pentaho Server, double click on the BATCH file. You have to wait until the Pentaho server is created.



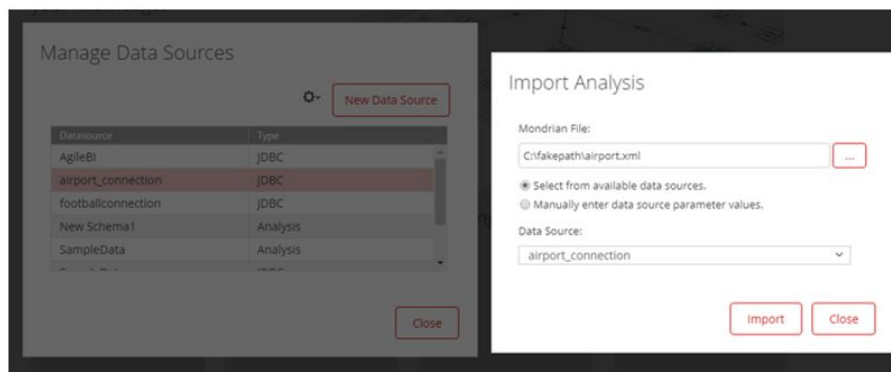
After waiting for the necessary time, proceed to connect to Pentaho Server from a browser by putting localhost: 8080 in the URL. The user is logged in: admin and password: password.



The connection to the *AIRPORT* database is created.



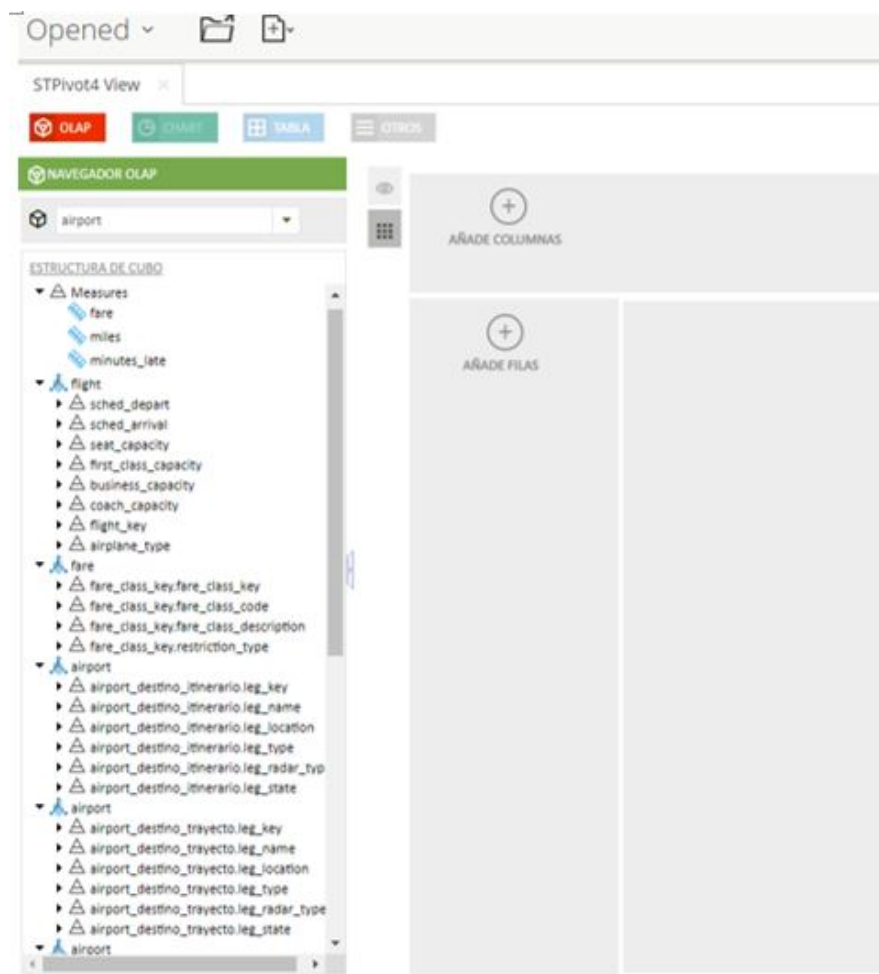
Next, the XML file (the cube created in Schema Workbench) is imported into Pentaho Server using the Manage Data Source (Import Analysis) tool.



Once the XML file is imported, we create STPivot on the cube to proceed with the analysis.



Finally, dragging both measures and dimensions (or their attributes) to columns or rows, you get, through a correct analysis, answers that can generate value to the company.



Questions asked:

1. What is the payment method and the place where you buy first class seats?

		channel_name				
		All channel.channel_names	Cash	Credit Card	Debit Card	PayPal
customer_location	first_class_capacity	Measures	Measures	Measures	Measures	Measures
		fare	fare	fare	fare	fare
- All customer.customer_locations	+ All flight.first_class_capacitys	54.142	10.666	13.371	16.231	13.874
+ California	+ All flight.first_class_capacitys	2.469	545	1.363	9	552
+ Wyoming	+ All flight.first_class_capacitys	2.132	736	299	597	500
+ Illinois	+ All flight.first_class_capacitys	2.035	500	911		624
+ Alabama	+ All flight.first_class_capacitys	2.009	361	760	465	423
+ Delaware	+ All flight.first_class_capacitys	1.973	264		879	830
+ Colorado	+ All flight.first_class_capacitys	1.902	144	565	515	678
+ Indiana	+ All flight.first_class_capacitys	1.782	662	104	872	144
+ Georgia	+ All flight.first_class_capacitys	1.679	302	519	683	175
+ South Carolina	+ All flight.first_class_capacitys	1.487	14		661	812
+ Kentucky	+ All flight.first_class_capacitys	1.438	274	280	530	354
+ North Dakota	+ All flight.first_class_capacitys	1.373	333	430	610	
+ Iowa	+ All flight.first_class_capacitys	1.327	15		730	582
+ New Mexico	+ All flight.first_class_capacitys	1.310	263	97	249	701
+ Florida	+ All flight.first_class_capacitys	1.298		510	313	475
+ South Dakota	+ All flight.first_class_capacitys	1.218	351	841	26	
+ Arizona	+ All flight.first_class_capacitys	1.218	345	277	359	237
+ New Hampshire	+ All flight.first_class_capacitys	1.138	343	245	338	212
+ Idaho	+ All flight.first_class_capacitys	1.121	452		587	82
+ Louisiana	+ All flight.first_class_capacitys	1.118		754		364
+ Virginia	+ All flight.first_class_capacitys	1.113	7		200	906

2. What type of aircraft, depending on the state, is the one with the most delay?

		Measures
airport_origen_itinerario.leg_location	airplane_type	minutes_late
- All airport.leg_locations	+ All flight.airplane_types	1.057
+ NC	- All flight.airplane_types	159
	Super 80	80
	727	79
	DC-10	0
+ MN	- All flight.airplane_types	132
	727	132
	Super 80	0
	DC-10	0
+ IL	- All flight.airplane_types	132
	727	132
	Super 80	0
	DC-10	0
+ MA	- All flight.airplane_types	96
	Super 80	96
	DC-10	0
	727	0
+ MO	- All flight.airplane_types	82
	727	82
	Super 80	0
	DC-10	0
+ OR	+ All flight.airplane_types	78
+ NY	+ All flight.airplane_types	77
+ TN	+ All flight.airplane_types	75
+ FL	+ All flight.airplane_types	72

3. What is the city, depending on the political state, where more money has been spent?

		Measures
airport_destino_itinerario.leg_location	customer_marital	fare
Chicago	— All customer.customer_maritals	9.416
	Married	5.854
	Single	2.912
	Coresident	650
San Diego	— All customer.customer_maritals	3.029
	Single	1.349
	Married	1.302
	Coresident	378
Costa Mesa	+ All customer.customer_maritals	1.583
Los Angeles	+ All customer.customer_maritals	1.341
New York	— All customer.customer_maritals	5.062
	Married	2.696
	Single	1.494
	Coresident	872
Washington	— All customer.customer_maritals	4.514
	Married	3.569
	Single	618
	Coresident	327
Seattle	— All customer.customer_maritals	4.229
	Married	2.162
	Single	1.481
	Coresident	586
Minneapolis	— All customer.customer_maritals	3.617
	Married	2.766

4. What type of plane, leaving Los Angeles, travels more miles?

		Measures
airport_origen_trayecto.leg_location	airplane_type	miles
– All airport.leg_locations	+ All flight.airplane_types	284.899
– CA	– All flight.airplane_types	41.553
	Super 80	18.483
	DC-10	13.100
	727	9.970
Los Angeles	– All flight.airplane_types	15.461
	DC-10	7.650
	Super 80	6.313
	727	1.498
San Diego	– All flight.airplane_types	13.428
	727	6.126
	DC-10	4.064
	Super 80	3.238
Costa Mesa	+ All flight.airplane_types	12.664
+ DC	+ All flight.airplane_types	24.900
+ NY	+ All flight.airplane_types	24.211
+ IL	+ All flight.airplane_types	22.494
+ FL	+ All flight.airplane_types	19.429
+ MO	+ All flight.airplane_types	18.959
+ OR	+ All flight.airplane_types	18.830
+ TN	+ All flight.airplane_types	17.531
+ TX	+ All flight.airplane_types	17.156
+ NC	+ All flight.airplane_types	16.517
+ MN	+ All flight.airplane_types	16.466