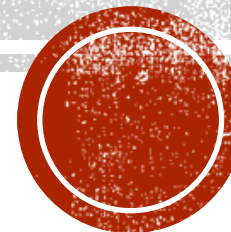


ANÁLISIS BIG DATA

Por Andrea Manuel Simón



ÍNDICE

- Información de los datos
- Cambio de str a int
- Análisis estadísticos:
 - Media
 - Desviación estándar
 - Matriz de correlación



INFORMACIÓN DE LOS DATOS

Data columns (total 16 columns):

#	Column	Non-Null Count	Dtype
0	Activity Period	15007 non-null	int64
1	Operating Airline	15007 non-null	object
2	Operating Airline IATA Code	14953 non-null	object
3	Published Airline	15007 non-null	object
4	Published Airline IATA Code	14953 non-null	object
5	GEO Summary	15007 non-null	object
6	GEO Region	15007 non-null	object
7	Activity Type Code	15007 non-null	object
8	Price Category Code	15007 non-null	object
9	Terminal	15007 non-null	object
10	Boarding Area	15007 non-null	object
11	Passenger Count	15007 non-null	int64
12	Adjusted Activity Type Code	15007 non-null	object
13	Adjusted Passenger Count	15007 non-null	int64
14	Year	15007 non-null	int64
15	Month	15007 non-null	object

dtypes: int64(4), object(12)

Información de las columnas

Operating Airline IATA Code

4T	312.625000
5Y	34.000000
9W	4280.312500
A8	5.000000
AA	127164.389706
...	
XE	5631.843750
XJ	2864.727273
XP	73.000000
YV	3710.581197
YX	3883.000000

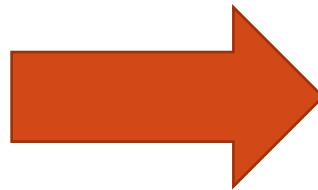
Media de los pasajeros en cada compañía

Name: Passenger Count, Length: 73, dtype: float64



CAMBIO DE STR A INT

- Price Category Code ['Low Fare' 'Other']
- Activity Type Code ['Deplaned' 'Enplaned' 'Thru / Transit']
- Terminal ['Terminal 1' 'International' 'Terminal 3' 'Other' 'Terminal 2']
- GEO Summary ['Domestic' 'International']
- Adjusted Activity Type Code ['Deplaned' 'Enplaned' 'Thru / Transit * 2']
- Month ['July' 'August' 'September' 'October' 'November' 'April' 'December' 'January' 'February' 'March' 'May' 'June']



- Price Category Code: 0 = low, 1 = other
- Geo Summary: 0 = Domestic, 1 = International
- Adjusted Activity Type Code: 0 = deplaned, 1 = Enplaned, 2 = Thru/Transit * 2
- Terminal: 0 = International, 1 = Terminal 1, 2 = Terminal 2, 3 = Terminal 3, 4 = Other
- A cada mes se le asignara el número real de dicho mes.



MEDIA

Activity Period 201045.073366
GEO Summary 0.613714
Price Category Code 0.872060
Terminal 1.000200
Passenger Count 29240.521090
Adjusted Activity Type Code 0.590125
Adjusted Passenger Count 29331.917105
Year 2010.385220
Month 6.551343
dtype: float64

	Activity Period	GEO Summary	Price Category Code	Terminal \
Boarding Area				
A	201074.682488	0.847656	0.879809	0.996172
B	200976.570998	0.108379	0.613648	0.000000
C	200992.599349	0.000000	0.820847	0.000000
D	201346.314815	0.185185	0.444444	4.000000
E	200917.158145	0.205707	0.970273	2.000000
F	201035.737110	0.445171	0.955701	2.000000
G	201064.539329	0.931112	0.990982	1.000000
Other	200725.740741	0.074074	1.000000	3.000000

	Passenger Count	Adjusted Activity Type Code \
Boarding Area		
A	11115.767656	0.505837
B	33804.871049	0.639739
C	34423.159609	0.542345
D	105124.197531	0.592593
E	48617.014269	0.689655
F	100600.343500	0.730574

GEO Region	Adjusted Passenger Count	Year	Month
Asia	13508.552704	2010.396578	6.535900
Australia / Oceania	6495.104478	2009.869742	6.483039
Canada	9803.791255	2010.635402	6.611425
Central America	4946.715328	2010.656934	6.583942
Europe	12779.055050	2010.673528	6.584969
Mexico	7250.898655	2010.588341	6.445740
Middle East	8658.612150	2012.560748	6.401869
South America	2786.011111	2011.866667	6.600000
US	58485.878385	2010.124030	6.565465

Operating Airline	Activity Period	GEO Summary	Price Category Code \
ATA Airlines	200586.363636	0.068182	0.0
Aer Lingus	201151.469388	1.000000	1.0
Aeromexico	201207.533333	1.000000	1.0
Air Berlin	201107.500000	1.000000	1.0
Air Canada	201123.497268	1.000000	1.0
...
Virgin Atlantic	201043.744186	1.000000	1.0
WestJet Airlines	201125.844660	1.000000	1.0
World Airways	201008.333333	0.666667	1.0
XL Airways France	201339.096774	1.000000	1.0
Xtra Airways	200608.000000	0.000000	1.0



DESVIACIÓN ESTÁNDAR

```
Activity Period      313.336196
GEO Summary         0.486914
Price Category Code 0.334034
Terminal            0.751869
Passenger Count     58319.509284
Adjusted Activity Type Code 0.603748
Adjusted Passenger Count 58284.182219
Year                3.137589
Month               3.464354
dtype: float64
```

	Activity Period	GEO Summary	Price Category Code	Terminal \
Boarding Area				
A	314.640882	0.359389	0.325216	0.061756
B	294.255947	0.310937	0.487035	0.000000
C	305.946668	0.000000	0.383637	0.000000
D	151.640665	0.389049	0.497673	0.000000
E	293.286164	0.404458	0.169933	0.000000
F	307.903263	0.497165	0.205834	0.000000
G	313.749343	0.253295	0.094546	0.000000
Other	187.786620	0.266880	0.000000	0.000000

	Passenger Count	Adjusted Activity Type Code \
Boarding Area		
A	13624.028630	0.524673
B	38938.939200	0.627088
C	40149.197576	0.549726
D	62710.950791	0.492112
E	71298.023217	0.676123
F	139056.322983	0.702489

GEO Region	Adjusted Passenger Count	Year	Month
Asia	16147.810667	3.141101	3.491609
Australia / Oceania	2650.383265	2.992176	3.481324
Canada	7805.730644	3.210175	3.414744
Central America	1220.840313	3.251234	3.544882
Europe	8602.128044	3.170094	3.399006
Mexico	5274.346847	3.179553	3.498072
Middle East	2732.719518	2.235764	3.625315
South America	396.758651	1.153402	3.556289
US	84859.991540	3.102956	3.465904

Operating Airline	Activity Period	GEO Summary	Price Category Code \
ATA Airlines	83.311992	0.254972	0.0
Aer Lingus	331.485075	0.000000	0.0
Aeromexico	218.109152	0.000000	0.0
Air Berlin	82.825979	0.000000	0.0
Air Canada	298.821335	0.000000	0.0
...
Virgin Atlantic	312.743907	0.000000	0.0
WestJet Airlines	299.812711	0.000000	0.0
World Airways	175.514482	0.577350	0.0
XL Airways France	113.622872	0.000000	0.0
Xtra Airways	0.000000	0.000000	0.0



MATRIZ DE CORRELACIÓN

	Activity Period	GEO Summary	Price Category Code	Terminal	Passenger Count	Adjusted Activity Type Code	Adjusted Passenger Count	Year	Month
Activity Period	1.000000	0.066100	-0.005754	0.103760	0.060311	-0.052450	0.059336	0.999940	-0.116571
GEO Summary	0.066100	1.000000	0.411498	0.136915	-0.395743	-0.026760	-0.396856	0.066046	-0.003234
Price Category Code	-0.005754	0.411498	1.000000	0.096685	-0.065047	0.001004	-0.064661	-0.005683	-0.005725
Terminal	0.103760	0.136915	0.096685	1.000000	0.260570	0.036001	0.261506	0.103611	0.000776
Passenger Count	0.060311	-0.395743	-0.065047	0.260570	1.000000	-0.071423	0.999941	0.060069	0.014521
Adjusted Activity Type Code	-0.052450	-0.026760	0.001004	0.036001	-0.071423	1.000000	-0.067804	-0.052364	-0.001360
Adjusted Passenger Count	0.059336	-0.396856	-0.064661	0.261506	0.999941	-0.067804	1.000000	0.059096	0.014503
Year	0.999940	0.066046	-0.005683	0.103611	0.060069	-0.052364	0.059096	1.000000	-0.127455
Month	-0.116571	-0.003234	-0.005725	0.000776	0.014521	-0.001360	0.014503	-0.127455	1.000000



FIN

