SweetViz
2.1.4
Get updates, docs & report issues here
Created & maintained by Francois Bertrand
Graphic design by Jean-Francois Hains

DataFrame   **NO COMPARISON TARGET**

327346	ROWS
0	DUPLICATES
108.0 MB	RAM
14	FEATURES
3	CATEGORICAL
9	NUMERICAL
2	TEXT

ASSOCIATIONS

Associations
[Only including
■ **Squares** are
values indicate

• **Circles** are the

DataFrame
Avg. arr_delay ●—

arr_delay

VALUES:	327,346 (100%)	MAX	1,272	RANGE	1,358	
MISSING:	---	95%	91	IQR	31.0	
		Q3	14	STD	44.6	
DISTINCT:	577 (<1%)	AVG	7	VAR	1,992	
		MEDIAN	-5			
ZEROES:	5,409 (2%)	Q1	-17	KURT.	29.2	
		5%	-32	SKEW	3.72	
		MIN	-86	SUM	2.3M	
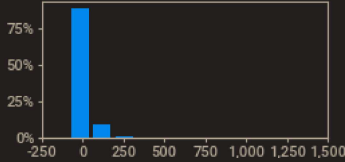
1 year

VALUES:	327,346 (100%)	
MISSING:	---	
DISTINCT:	1 (<1%)	
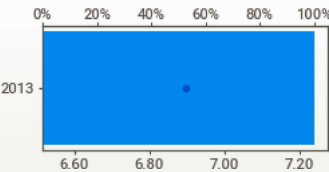
2 month

VALUES:	327,346 (100%)	MAX	12.0	RANGE	11.0
MISSING:	---	95%	12.0	IQR	6.00
		Q3	10.0	STD	3.41
DISTINCT:	12 (<1%)	MEDIAN	7.0	VAR	11.7
		AVG	6.6		
ZEROES:	---	Q1	4.0	KURT.	-1.19
		5%	1.0	SKEW	-0.024
		MIN	1.0	SUM	2.1M
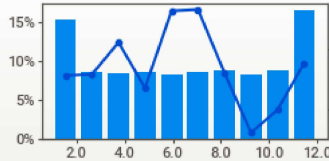
3 day

VALUES:	327,346 (100%)	MAX	31.0	RANGE	30.0
MISSING:	---	95%	29.0	IQR	15.0
		Q3	23.0	STD	8.78
DISTINCT:	31 (<1%)	MEDIAN	16.0	VAR	77.0
		AVG	15.7		
ZEROES:	---	Q1	8.0	KURT.	-1.19
		5%	2.0	SKEW	-0.001
		MIN	1.0	SUM	5.2M
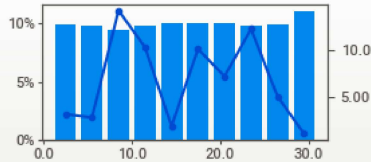
4 dep_delay

VALUES:	327,346 (100%)	MAX	1,301	RANGE	1,344
MISSING:	---	95%	88	IQR	16.0
		Q3	11	STD	40.1
DISTINCT:	526 (<1%)	AVG	13	VAR	1,605
		MEDIAN	-2		
ZEROES:	16,466 (5%)	Q1	-5	KURT.	44.4
		5%	-9	SKEW	4.82
		MIN	-43	SUM	4.1M

5 carrier

VALUES:	327,346 (100%)	
MISSING:	---	
DISTINCT:	16 (<1%)	

UA
B6
EV
DL
(Other)
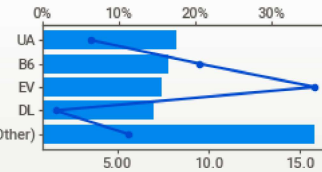
6 flight

VALUES:	327,346 (100%)	MAX	8,500	RANGE	8,499
MISSING:	---	95%	4,689	IQR	2,868
		Q3	3,412	STD	1,622
DISTINCT:	3,835 (1%)	AVG	1,943	VAR	2.6M
		MEDIAN	1,467		
ZEROES:	---	Q1	544	KURT.	-0.791
		5%	87	SKEW	0.693
		MIN	1	SUM	636.1M
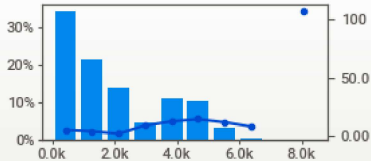
7 tailnum

VALUES:	327,346 (100%)	
MISSING:	---	
DISTINCT:	4,037 (1%)	

544	<1%	N725MQ
485	<1%	N722MQ
475	<1%	N723MQ
462	<1%	N711MQ
449	<1%	N713MQ
428	<1%	N258JB
403	<1%	N353JB
324,108	>99%	(Other)

8 ⊞ origin

VALUES:	327,346 (100%)	
MISSING:	---	
DISTINCT:	3 (<1%)	

9 ▤ dest

VALUES:	327,346 (100%)	
MISSING:	---	
DISTINCT:	104 (<1%)	

16,837	5%	ATL
16,566	5%	ORD
16,026	5%	LAX
15,022	5%	BOS
13,967	4%	MCO
13,674	4%	CLT
13,173	4%	SFO
222,081	68%	(Other)

10 ⌁ air_time

VALUES:	327,346 (100%)	
MISSING:	---	
DISTINCT:	509 (<1%)	
ZEROES:	---	

MAX	695		RANGE	675
95%	339		IQR	110
Q3	192		STD	93.7
AVG	151		VAR	8,777
MEDIAN	129			
Q1	82		KURT.	0.863
5%	40		SKEW	1.07
MIN	20		SUM	49.3M

11 ⌁ distance

VALUES:	327,346 (100%)	
MISSING:	---	
DISTINCT:	213 (<1%)	
ZEROES:	---	

MAX	4,983		RANGE	4,903
95%	2,475		IQR	880
Q3	1,389		STD	736
AVG	1,048		VAR	542k
MEDIAN	888			
Q1	509		KURT.	1.15
5%	199		SKEW	1.11
MIN	80		SUM	343.2M

12 ⌁ hour

VALUES:	327,346 (100%)	
MISSING:	---	
DISTINCT:	19 (<1%)	
ZEROES:	---	

MAX	23.0		RANGE	18.0
95%	20.0		IQR	8.00
Q3	17.0		STD	4.66
AVG	13.1		VAR	21.7
MEDIAN	13.0			
Q1	9.0		KURT.	-1.21
5%	6.0		SKEW	0.012
MIN	5.0		SUM	4.3M

13 ⌁ minute

VALUES:	327,346 (100%)	
MISSING:	---	
DISTINCT:	60 (<1%)	
ZEROES:	58,924 (18%)	

MAX	59.0		RANGE	59.0
95%	58.0		IQR	36.0
Q3	44.0		STD	19.3
MEDIAN	29.0		VAR	372
AVG	26.2			
Q1	8.0		KURT.	-1.23
5%	0.0		SKEW	0.093
MIN	0.0		SUM	8.6M

DataFrame   NO COMPARISON TARGET

327346	ROWS	
0	DUPLICATES	
108.0 MB	RAM	
14	FEATURES	
3	CATEGORICAL	
9	NUMERICAL	
2	TEXT	

ASSOCIATIONS

DataFrame
Avg. arr_delay

MAX	1,272	RANGE	1,358
5%	91	IQR	31.0
3	14	STD	44.6
AVG	7	VAR	1,992
MEDIAN	-5		
1	-17	KURT.	29.2
%	-32	SKEW	3.72
MIN	-86	SUM	2.3M

2013

MAX	12.0	RANGE	11.0
5%	12.0	IQR	6.00
3	10.0	STD	3.41
MEDIAN	7.0	VAR	11.7
AVG	6.6		
1	4.0	KURT.	-1.19
%	1.0	SKEW	-0.024
MIN	1.0	SUM	2.1M

MAX	31.0	RANGE	30.0
5%	29.0	IQR	15.0
3	23.0	STD	8.78
MEDIAN	16.0	VAR	77.0
AVG	15.7		

Associations

[Only including dataset "DataFrame"]
■ **Squares** are categorical associations (uncertainty coefficient & correlation ratio) from 0 to 1. The uncertainty coefficient is **assymmetrical**, (i.e. ROW LABEL values indicate how much they PROVIDE INFORMATION to each LABEL at the TOP).

• **Circles** are the symmetrical numerical correlations (Pearson's) from -1 to 1. The **trivial diagonal** is intentionally left blank for clarity.

Row/column labels: arr_delay, year, month, day, dep_delay, carrier, flight, origin, air_time, distance, hour, minute