

Project 2 -- Seyi Ogunmodede

**Instructor Help: Dr. Mark Daniel Ward

- Help with figuring out how to write a function.
- Help to navigate through the instructions
- Got help through the video

Collaboration:

Question 1

In [5]:

```
%%R  
myDF <- read.csv("/anvil/projects/tdm/data/flights/subset/1995.csv",  
                  stringsAsFactors = TRUE)
```

In [6]:

```
%%R  
head(myDF)
```

	Year	Month	DayofMonth	DayOfWeek	DepTime	CRSDepTime	ArrTime	CRSArrTime
1	1995	1	6	5	657	645	952	937
2	1995	1	7	6	648	645	938	937
3	1995	1	8	7	649	645	932	937
4	1995	1	9	1	645	645	928	937
5	1995	1	10	2	645	645	931	937
6	1995	1	11	3	646	645	929	937
	UniqueCarrier	FlightNum	TailNum	ActualElapsedTime	CRSElapsedTime	AirTime		
1	UA	482	N7298U		115		112	83
2	UA	482	N7449U		110		112	88
3	UA	482	N7453U		103		112	83
4	UA	482	N7288U		103		112	84
5	UA	482	N7275U		106		112	82
6	UA	482	N7281U		103		112	85
	ArrDelay	DepDelay	Origin	Dest	Distance	TaxiIn	TaxiOut	Cancelled
1	15	12	ORD	PHL	678	7	25	0
2	1	3	ORD	PHL	678	5	17	0
3	-5	4	ORD	PHL	678	3	17	0
4	-9	0	ORD	PHL	678	3	16	0
5	-6	0	ORD	PHL	678	6	18	0
6	-8	1	ORD	PHL	678	5	13	0
	CancellationCode	Diverted	CarrierDelay	WeatherDelay	NASDelay	SecurityDelay		
1	NA	0	NA	NA	NA	NA	NA	NA
2	NA	0	NA	NA	NA	NA	NA	NA
3	NA	0	NA	NA	NA	NA	NA	NA
4	NA	0	NA	NA	NA	NA	NA	NA
5	NA	0	NA	NA	NA	NA	NA	NA
6	NA	0	NA	NA	NA	NA	NA	NA
	LateAircraftDelay							
1		NA						
2		NA						
3		NA						
4		NA						
5		NA						
6		NA						

In [4]: `%%R
dim(myDF)`

[1] 5327435 29

In [6]: `%%R
str(myDF)`

```
'data.frame': 5327435 obs. of 29 variables:  
 $ Year           : int 1995 1995 1995 1995 1995 1995 1995 1995 1995 1995 ...  
 $ Month          : int 1 1 1 1 1 1 1 1 1 1 ...  
 $ DayofMonth     : int 6 7 8 9 10 11 12 13 14 15 ...  
 $ DayOfWeek      : int 5 6 7 1 2 3 4 5 6 7 ...  
 $ DepTime         : int 657 648 649 645 645 646 NA 644 644 643 ...  
 $ CRSDepTime    : int 645 645 645 645 645 645 645 645 645 645 ...  
 $ ArrTime         : int 952 938 932 928 931 929 NA 953 938 940 ...  
 $ CRSArrTime    : int 937 937 937 937 937 937 937 937 937 937 ...  
 $ UniqueCarrier   : Factor w/ 10 levels "AA","AS","CO",...: 8 8 8 8 8 8 8 8 8 8 ...  
 $ FlightNum       : int 482 482 482 482 482 482 482 482 482 482 ...  
 $ TailNum         : Factor w/ 3508 levels "112","EI-BOM",...: 2526 2561 2566 2511 24  
 89 2498 3508 2467 2504 2484 ...  
 $ ActualElapsedTime: int 115 110 103 103 106 103 NA 129 114 117 ...  
 $ CRSElapsedTime  : int 112 112 112 112 112 112 112 112 112 112 ...  
 $ AirTime          : int 83 88 83 84 82 85 45 110 94 99 ...  
 $ ArrDelay         : int 15 1 -5 -9 -6 -8 NA 16 1 3 ...  
 $ DepDelay         : int 12 3 4 0 0 1 NA -1 -1 -2 ...  
 $ Origin           : Factor w/ 218 levels "ABE","ABQ","ACK",...: 157 157 157 157 157  
 157 157 157 157 157 ...  
 $ Dest              : Factor w/ 218 levels "ABE","ABQ","ACK",...: 165 165 165 165 165  
 165 165 165 165 165 ...  
 $ Distance          : int 678 678 678 678 678 678 678 678 678 678 ...  
 $ TaxiIn            : int 7 5 3 3 6 5 6 5 5 3 ...  
 $ TaxiOut           : int 25 17 17 16 18 13 10 14 15 15 ...  
 $ Cancelled         : int 0 0 0 0 0 0 1 0 0 0 ...  
 $ CancellationCode : logi NA NA NA NA NA NA ...  
 $ Diverted          : int 0 0 0 0 0 0 0 0 0 0 ...  
 $ CarrierDelay      : logi NA NA NA NA NA NA ...  
 $ WeatherDelay      : logi NA NA NA NA NA NA ...  
 $ NASDelay          : logi NA NA NA NA NA NA ...  
 $ SecurityDelay     : logi NA NA NA NA NA NA ...  
 $ LateAircraftDelay: logi NA NA NA NA NA NA ...
```

In [7]:

```
%%R  
tail(myDF)
```

	Year	Month	DayofMonth	DayOfWeek	DeptTime	CRSDepTime	ArrTime	CRSArrTime	
5327430	1995	12	11	1	2220	2215	2337	2332	
5327431	1995	12	12	2	2217	2215	2332	2332	
5327432	1995	12	13	3	2216	2215	2355	2332	
5327433	1995	12	14	4	2215	2215	2341	2332	
5327434	1995	12	15	5	2236	2215	2348	2332	
5327435	1995	12	17	7	2215	2215	2328	2332	
	UniqueCarrier	FlightNum	TailNum	ActualElapsedTime	CRSElapsedTime				
5327430	TW	427	N959U		137		137		
5327431	TW	427	N9409F		135		137		
5327432	TW	427	N912TW		159		137		
5327433	TW	427	N902TW		146		137		
5327434	TW	427	N954U		132		137		
5327435	TW	427	N920TW		133		137		
	AirTime	ArrDelay	DepDelay	Origin	Dest	Distance	TaxiIn	TaxiOut	Cancelled
5327430	122	5	5	STL	DEN	770	6	9	0
5327431	118	0	2	STL	DEN	770	5	12	0
5327432	129	23	1	STL	DEN	770	6	24	0
5327433	125	9	0	STL	DEN	770	5	16	0
5327434	114	16	21	STL	DEN	770	5	13	0
5327435	111	-4	0	STL	DEN	770	6	16	0
	CancellationCode	Diverted	CarrierDelay	WeatherDelay	NASDelay				
5327430	NA	0	NA	NA	NA				
5327431	NA	0	NA	NA	NA				
5327432	NA	0	NA	NA	NA				
5327433	NA	0	NA	NA	NA				
5327434	NA	0	NA	NA	NA				
5327435	NA	0	NA	NA	NA				
	SecurityDelay	LateAircraftDelay							
5327430	NA	NA							
5327431	NA	NA							
5327432	NA	NA							
5327433	NA	NA							
5327434	NA	NA							
5327435	NA	NA							

```
In [15]: %%R
class(myDF$Dest)

[1] "factor"
```

```
In [8]: %%R
class(myDF)

[1] "data.frame"
```

```
In [9]: %%R
summary(myDF$ArrDelay)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
-972.00	-6.00	1.00	7.02	11.00	1348.00	102397

```
In [10]: %%R
class(myDF$ArrDelay)

[1] "integer"
```

```
In [24]: %%R
dim(myDF)

[1] 5327435    29
```

In []:

In []:

In []:

A. It has 29 column
 B. It has 5327435 rows
 C. The type of data in the data frame are integer, factor and logical

Question 2

In [11]:

```
%%R
myairports <- myDF$Origin
```

In [12]:

```
%%R
str(myairports <- myDF$Origin)
```

```
Factor w/ 218 levels "ABE","ABQ","ACK",...: 157 157 157 157 157 157 157 157 157 157
...
...
```

In [29]:

```
%%R
length(myairports <- myDF$Origin)
```

```
[1] 5327435
```

In [26]:

```
%%R
class(myairports)
```

```
[1] "factor"
```

In [28]:

```
%%R
```

```
levels(myairports)
```

```
[1] "ABE" "ABQ" "ACK" "ADQ" "AGS" "AKN" "ALB" "AMA" "ANC" "ANI" "ATL" "AUS"
[13] "AVL" "AVP" "AZO" "BDL" "BET" "BFL" "BGM" "BGR" "BHM" "BIL" "BIS" "BNA"
[25] "BOI" "BOS" "BQN" "BRO" "BRW" "BTR" "BTW" "BUF" "BUR" "BWI" "BNZ" "CAE"
[37] "CAK" "CDV" "CHA" "CHS" "CID" "CLE" "CLT" "CMH" "COS" "CRP" "CRW" "CSG"
[49] "CVG" "DAB" "DAL" "DAY" "DCA" "DEN" "DFW" "DLG" "DLH" "DRO" "DSM" "DTW"
[61] "DUT" "EGE" "ELM" "ELP" "ERI" "EUG" "EVV" "EWR" "FAI" "FAR" "FAT" "FAY"
[73] "FCA" "FLL" "FNT" "FSD" "FWA" "GEG" "GFK" "GNV" "GRB" "GRR" "GSO" "GSP"
[85] "GST" "GTF" "GUC" "HDN" "HLN" "HNL" "HOU" "HPN" "HRL" "HSV" "HTS" "HVN"
[97] "IAD" "IAH" "ICT" "IDA" "ILM" "IND" "ISP" "ITH" "JAC" "JAN" "JAX" "JFK"
[109] "JNU" "KOA" "KSM" "KTN" "LAN" "LAS" "LAX" "LBB" "LEX" "LFT" "LGA" "LGB"
[121] "LIT" "LNK" "LSE" "LWB" "MAF" "MBS" "MCI" "MCO" "MDT" "MDW" "MEM" "MFE"
[133] "MFR" "MGM" "MHT" "MIA" "MKE" "MLB" "MLI" "MLU" "MOB" "MOT" "MRY" "MSN"
[145] "MSO" "MSP" "MSY" "MTJ" "MYR" "OAJ" "OAK" "OGG" "OKC" "OMA" "OME" "ONT"
[157] "ORD" "ORF" "OTZ" "PBI" "PDX" "PFN" "PHF" "PHL" "PHX" "PIA" "PIT" "PNS"
[169] "PSC" "PSG" "PSP" "PVD" "PWM" "RAP" "RDU" "RIC" "RNO" "ROA" "ROC" "RST"
[181] "RSW" "SAN" "SAT" "SAV" "SBA" "SBN" "SCC" "SDF" "SEA" "SFO" "SGF" "SHV"
[193] "SIT" "SJC" "SJU" "SLC" "SMF" "SNA" "SPN" "SRQ" "STL" "STT" "STX" "SUX"
[205] "SWF" "SYR" "TLH" "TOL" "TPA" "TRI" "TUL" "TUS" "TVC" "TWF" "TYS" "VPS"
[217] "WRG" "YAK"
```

In [31]:

```
%%R
```

```
sum(myDF$Origin=="ORD")
```

```
[1] 293422
```

In [14]:

```
%%R
length(myDF$Origin[myDF$Origin=="ORD"])
```

```
[1] 293422
```

In [16]:

```
%%R
head(myDF$Origin,n=250)
```

```
[1] ORD ORD
[19] ORD ORD
[37] ORD ORD
[55] ORD ORD
[73] ORD ORD
[91] ORD ORD
[109] ORD ORD
[127] ORD ORD
[145] ORD ORD
[163] ORD ORD
[181] ORD ORD
[199] ORD ORD
[217] ORD ORD
[235] ORD ORD ORD ORF ORF
218 Levels: ABE ABQ ACK ADQ AGS AKN ALB AMA ANC ANI ATL AUS AVL AVP AZO ... YAK
```

In [17]:

```
%%R
table(head(myDF$Origin,n=250))
```

	ABE	ABQ	ACK	ADQ	AGS	AKN	ALB	AMA	ANC	ANI	ATL	AUS	AVL	AVP	AZO	BDL	BET	BFL	BGM	BGR
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BHM	BIL	BIS	BNA	BOI	BOS	BQN	BRO	BRW	BTR	BTV	BUF	BUR	BWI	BZN	CAE	CAK	CDV	CHA	CHS	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
CID	CLE	CLT	CMH	COS	CRP	CRW	CSG	CVG	DAB	DAL	DAY	DCA	DEN	DFW	DLG	DLH	DRO	DSM	DTW	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
DUT	EGL	ELM	ELP	ERI	EUG	EVV	EWR	FAI	FAR	FAT	FAY	FCA	FLL	FNT	FSD	FWA	GEG	GFK	GNV	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
GRB	GRR	GSO	GSP	GST	GTF	GUC	HDN	HLN	HNL	HOU	HPN	HRL	HSV	HTS	HVN	IAD	IAH	ICT	IDA	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
ILM	IND	ISP	ITH	JAC	JAN	JAX	JFK	JNU	KOA	KSM	KTN	LAN	LAS	LAX	LBB	LEX	LFT	LGA	LGB	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
LIT	LNK	LSE	LWB	MAF	MBS	MCI	MCO	MDT	MDW	MEM	MFE	MFR	MGM	MHT	MIA	MKE	MLB	MLI	MLU	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
MOB	MOT	MRY	MSN	MSO	MSP	MSY	MTJ	MYR	OAJ	OAK	OGG	OKC	OMA	OME	ONT	ORD	ORF	OTZ	PBI	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	237	13	0	
PDX	PFN	PHF	PHL	PHX	PIA	PIT	PNS	PSC	PSG	PSP	PVD	PWM	RAP	RDU	RIC	RNO	ROA	ROC	RST	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
RSW	SAN	SAT	SAV	SBA	SBN	SCC	SDF	SEA	SFO	SGF	SHV	SIT	SJC	SJU	SLC	SMF	SNA	SPN	SRQ	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
STL	STT	STX	SUX	SWF	SYR	TLH	TOL	TPA	TRI	TUL	TUS	TVC	TWF	TYS	VPS	WRG	YAK			
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

In [18]:

```
%%R
table(head(myDF$Origin,n=250))["ORD"]
```

```
ORD
```

```
237
```

In [19]:

```
%%R
table(myDF$Origin)
```

ABE	ABQ	ACK	ADQ	AGS	AKN	ALB	AMA	ANC	ANI	ATL
7198	36871	40	504	2387	475	11997	4294	19040	133	230470
AUS	AVL	AVP	AZO	BDL	BET	BFL	BGM	BGR	BHM	BIL
36474	2877	2605	2490	25771	981	455	1422	1589	18118	3701
BIS	BNA	BOI	BOS	BQN	BRO	BRW	BTR	BTV	BUF	BUR
1417	43382	10174	92852	734	687	1024	4741	3362	19932	31472
BWI	BZN	CAE	CAK	CDV	CHA	CHS	CID	CLE	CLT	CMH
66154	2205	5714	1574	726	3387	9228	5318	66888	131417	38271
COS	CRP	CRW	CSG	CVG	DAB	DAL	DAY	DCA	DEN	DFW
12077	4225	2031	971	76629	4364	46570	14524	86266	130327	268127
DLG	DLH	DRO	DSM	DTW	DUT	EGE	ELM	ELP	ERI	EUG
367	1813	102	9177	147610	379	690	1392	28728	1454	2173
EVV	EWR	FAI	FAR	FAT	FAY	FCA	FLL	FNT	FSD	FWA
709	109592	4843	2339	2507	2188	1018	33641	248	4343	3354
GEG	GFK	GNV	GRB	GRR	GSO	GSP	GST	GTF	GUC	HDN
12230	1414	1473	2843	9171	32676	8835	98	2428	243	556
HLN	HNL	HOU	HPN	HRL	HSV	HTS	HVN	IAD	IAH	ICT
1094	12537	57492	6228	5752	6957	615	1081	33053	121029	7051
IDA	ILM	IND	ISP	ITH	JAC	JAN	JAX	JFK	JNU	KOA
1298	2204	37092	3111	1056	1424	4945	18184	46187	4130	455
KSM	KTN	LAN	LAS	LAX	LBB	LEX	LFT	LGA	LGB	LIT
133	2589	2365	112181	179909	6338	5070	2	99225	1401	16652
LNK	LSE	LWB	MAF	MBS	MCI	MCO	MDT	MDW	MEM	MFE
3488	1392	4	7512	3861	56727	67900	8648	47985	53388	4057
MFR	MGM	MHT	MIA	MKE	MLB	MLI	MLU	MOB	MOT	MRY
1342	2498	4629	70581	17172	3171	1641	2209	5378	1091	1205
MSN	MSO	MSP	MSY	MTJ	MYR	OAJ	OAK	OGG	OKC	OMA
4640	1822	136436	49179	45	2720	369	65541	2609	22077	18215
OME	ONT	ORD	ORF	OTZ	PBI	PDX	PFN	PHF	PHL	PHX
1013	40062	293422	18198	1022	22558	48049	1320	1016	91577	157155
PIA	PIT	PNS	PSC	PSG	PSP	PVD	PWM	RAP	RDU	RIC
76	126190	6621	1454	725	3132	12530	5451	1077	27642	15990
RNO	ROA	ROC	RST	RSW	SAN	SAT	SAV	SBA	SBN	SCC
20727	3312	16130	3514	14990	65731	38060	5763	1480	3560	632
SDF	SEA	SFO	SGF	SHV	SIT	SJC	SJU	SLC	SMF	SNA
24094	92016	129137	1575	4552	1311	39283	21458	84056	36051	36159
SPN	SRQ	STL	STT	STX	SUX	SWF	SYR	TLH	TOL	TPA
105	6863	174847	2874	1059	710	4402	12592	4668	2461	45782
TRI	TUL	TUS	TVC	TWF	TYS	VPS	WRG	YAK		
3056	20795	18549	1055	23	9509	1094	726	726		

In [20]:

```
%%R
sort(table(myDF$Origin))
```

project 2

LFT	LWB	TWF	ACK	MTJ	PIA	GST	DRO	SPN	ANI	KSM
2	4	23	40	45	76	98	102	105	133	133
GUC	FNT	DLG	OAJ	DUT	BFL	KOA	AKN	ADQ	HDN	HTS
243	248	367	369	379	455	455	475	504	556	615
SCC	BRO	EGE	EVV	SUX	PSG	CDV	WRG	YAK	BQN	CSG
632	687	690	709	710	725	726	726	726	734	971
BET	OME	PHF	FCA	OTZ	BRW	TVC	ITH	STX	RAP	HVN
981	1013	1016	1018	1022	1024	1055	1056	1059	1077	1081
MOT	HLN	VPS	MRY	IDA	SIT	PFN	MFR	ELM	LSE	LGB
1091	1094	1094	1205	1298	1311	1320	1342	1392	1392	1401
GFK	BIS	BGM	JAC	ERI	PSC	GNV	SBA	CAK	SGF	BGR
1414	1417	1422	1424	1454	1454	1473	1480	1574	1575	1589
MLI	DLH	MSO	CRW	EUG	FAY	ILM	BZN	MLU	FAR	LAN
1641	1813	1822	2031	2173	2188	2204	2205	2209	2339	2365
AGS	GTF	TOL	AZO	MGM	FAT	KTN	AVP	OGG	MYR	GRB
2387	2428	2461	2490	2498	2507	2589	2605	2609	2720	2843
STT	AVL	TRI	ISP	PSP	MLB	ROA	FWA	BTV	CHA	LNK
2874	2877	3056	3111	3132	3171	3312	3354	3362	3387	3488
RST	SBN	BIL	MBS	MFE	JNU	CRP	AMA	FSD	DAB	SWF
3514	3560	3701	3861	4057	4130	4225	4294	4343	4364	4402
SHV	MHT	MSN	TLH	BTR	FAI	JAN	LEX	CID	MOB	PWM
4552	4629	4640	4668	4741	4843	4945	5070	5318	5378	5451
CAE	HRL	SAV	HPN	LBB	PNS	SRQ	HSV	ICT	ABE	MAF
5714	5752	5763	6228	6338	6621	6863	6957	7051	7198	7512
MDT	GSP	GRR	DSM	CHS	TYS	BOI	ALB	COS	GEG	PVD
8648	8835	9171	9177	9228	9509	10174	11997	12077	12230	12530
HNL	SYR	DAY	RSW	RIC	ROC	LIT	MKE	BHM	JAX	ORF
12537	12592	14524	14990	15990	16130	16652	17172	18118	18184	18198
OMA	TUS	ANC	BUF	RNO	TUL	SJU	OKC	PBI	SDF	BDL
18215	18549	19040	19932	20727	20795	21458	22077	22558	24094	25771
RDU	ELP	BUR	GSO	IAD	FLL	SMF	SNA	AUS	ABQ	IND
27642	28728	31472	32676	33053	33641	36051	36159	36474	36871	37092
SAT	CMH	SJC	ONT	BNA	TPA	JFK	DAL	MDW	PDX	MSY
38060	38271	39283	40062	43382	45782	46187	46570	47985	48049	49179
MEM	MCI	HOU	OAK	SAN	BWI	CLE	MCO	MIA	CVG	SLC
53388	56727	57492	65541	65731	66154	66888	67900	70581	76629	84056
DCA	PHL	SEA	BOS	LGA	EWR	LAS	IAH	PIT	SFO	DEN
86266	91577	92016	92852	99225	109592	112181	121029	126190	129137	130327
CLT	MSP	DTW	PHX	STL	LAX	ATL	DFW	ORD		
131417	136436	147610	157155	174847	179909	230470	268127	293422		

```
In [12]: %>%  
tail(sort(table(myDF$Origin)))
```

PHX	STL	LAX	ATL	DFW	ORD
157155	174847	179909	230470	268127	293422

```
In [13]: %>%  
sort(table(myDF$Origin), decreasing=TRUE)
```

	ORD	DFW	ATL	LAX	STL	PHX	DTW	MSP	CLT	DEN	SFO
293422	268127	230470	179909	174847	157155	147610	136436	131417	130327	129137	
PIT	IAH	LAS	EWR	LGA	BOS	SEA	PHL	DCA	SLC	CVG	
126190	121029	112181	109592	99225	92852	92016	91577	86266	84056	76629	
MIA	MCO	CLE	BWI	SAN	OAK	HOU	MCI	MEM	MSY	PDX	
70581	67900	66888	66154	65731	65541	57492	56727	53388	49179	48049	
MDW	DAL	JFK	TPA	BNA	ONT	SJC	CMH	SAT	IND	ABQ	
47985	46570	46187	45782	43382	40062	39283	38271	38060	37092	36871	
AUS	SNA	SMF	FLL	IAD	GSO	BUR	ELP	RDU	BDL	SDF	
36474	36159	36051	33641	33053	32676	31472	28728	27642	25771	24094	
PBI	OKC	SJU	TUL	RNO	BUF	ANC	TUS	OMA	ORF	JAX	
22558	22077	21458	20795	20727	19932	19040	18549	18215	18198	18184	
BHM	MKE	LIT	ROC	RIC	RSW	DAY	SYR	HNL	PVD	GEG	
18118	17172	16652	16130	15990	14990	14524	12592	12537	12530	12230	
COS	ALB	BOI	TYC	CHS	DSM	GRR	GSP	MDT	MAF	ABE	
12077	11997	10174	9509	9228	9177	9171	8835	8648	7512	7198	
ICT	HSV	SRQ	PNS	LBB	HPN	SAV	HRL	CAE	PWM	MOB	
7051	6957	6863	6621	6338	6228	5763	5752	5714	5451	5378	
CID	LEX	JAN	FAI	BTR	TLH	MSN	MHT	SHV	SWF	DAB	
5318	5070	4945	4843	4741	4668	4640	4629	4552	4402	4364	
FSD	AMA	CRP	JNU	MFE	MBS	BIL	SBN	RST	LNK	CHA	
4343	4294	4225	4130	4057	3861	3701	3560	3514	3488	3387	
BTV	FWA	ROA	MLB	PSP	ISP	TRI	AVL	STT	GRB	MYR	
3362	3354	3312	3171	3132	3111	3056	2877	2874	2843	2720	
OGG	AVP	KTN	FAT	MGM	AZO	TOL	GTF	AGS	LAN	FAR	
2609	2605	2589	2507	2498	2490	2461	2428	2387	2365	2339	
MLU	BZN	ILM	FAY	EUG	CRW	MSO	DLH	MLI	BGR	SGF	
2209	2205	2204	2188	2173	2031	1822	1813	1641	1589	1575	
CAK	SBA	GNV	ERI	PSC	JAC	BGM	BIS	GFK	LGB	ELM	
1574	1480	1473	1454	1454	1424	1422	1417	1414	1401	1392	
LSE	MFR	PFN	SIT	IDA	MRY	HLN	VPS	MOT	HVN	RAP	
1392	1342	1320	1311	1298	1205	1094	1094	1091	1081	1077	
STX	ITH	TVC	BRW	OTZ	FCA	PHF	OME	BET	CSG	BQN	
1059	1056	1055	1024	1022	1018	1016	1013	981	971	734	
CDV	WRG	YAK	PSG	SUX	EVV	EGE	BRO	SCC	HTS	HDN	
726	726	726	725	710	709	690	687	632	615	556	
ADQ	AKN	BFL	KOA	DUT	OAJ	DLG	FNT	GUC	ANI	KSM	
504	475	455	455	379	369	367	248	243	133	133	
SPN	DRO	GST	PIA	MTJ	ACK	TWF	LWB	LFT			
105	102	98	76	45	40	23	4	2			

```
In [14]: %%R
head(sort(table(myDF$Origin), decreasing=TRUE))
```

```
ORD DFW ATL LAX STL PHX
293422 268127 230470 179909 174847 157155
```

- a. The type of data in the vector myairports is Factor with 218 levels. b. How many of the first 250 flights departed from O'Hare? 237. To get this you can use either the code with "ORD" or without. c. How many flights departed by O'Hare altogether in 1995? 293422. To get this you can use either of the last four code.

Question 3

```
In [15]: %%R
table(myDF$Origin)
```

ABE	ABQ	ACK	ADQ	AGS	AKN	ALB	AMA	ANC	ANI	ATL
7198	36871	40	504	2387	475	11997	4294	19040	133	230470
AUS	AVL	AVP	AZO	BDL	BET	BFL	BGM	BGR	BHM	BIL
36474	2877	2605	2490	25771	981	455	1422	1589	18118	3701
BIS	BNA	BOI	BOS	BQN	BRO	BRW	BTR	BTW	BUF	BUR
1417	43382	10174	92852	734	687	1024	4741	3362	19932	31472
BWI	BZN	CAE	CAK	CDV	CHA	CHS	CID	CLE	CLT	CMH
66154	2205	5714	1574	726	3387	9228	5318	66888	131417	38271
COS	CRP	CRW	CSG	CVG	DAB	DAL	DAY	DCA	DEN	DFW
12077	4225	2031	971	76629	4364	46570	14524	86266	130327	268127
DLG	DLH	DRO	DSM	DTW	DUT	EGE	ELM	ELP	ERI	EUG
367	1813	102	9177	147610	379	690	1392	28728	1454	2173
EVV	EWR	FAI	FAR	FAT	FAY	FCA	FLL	FNT	FSD	FWA
709	109592	4843	2339	2507	2188	1018	33641	248	4343	3354
GEG	GFK	GNV	GRB	GRR	GSO	GSP	GST	GTF	GUC	HDN
12230	1414	1473	2843	9171	32676	8835	98	2428	243	556
HLN	HNL	HOU	HPN	HRL	HSV	HTS	HVN	IAD	IAH	ICT
1094	12537	57492	6228	5752	6957	615	1081	33053	121029	7051
IDA	ILM	IND	ISP	ITH	JAC	JAN	JAX	JFK	JNU	KOA
1298	2204	37092	3111	1056	1424	4945	18184	46187	4130	455
KSM	KTN	LAN	LAS	LAX	LBB	LEX	LFT	LGA	LGB	LIT
133	2589	2365	112181	179909	6338	5070	2	99225	1401	16652
LNK	LSE	LWB	MAF	MBS	MCI	MCO	MDT	MDW	MEM	MFE
3488	1392	4	7512	3861	56727	67900	8648	47985	53388	4057
MFR	MGM	MHT	MIA	MKE	MLB	MLI	MLU	MOB	MOT	MRY
1342	2498	4629	70581	17172	3171	1641	2209	5378	1091	1205
MSN	MSO	MSP	MSY	MTJ	MYR	OAJ	OAK	OGG	OKC	OMA
4640	1822	136436	49179	45	2720	369	65541	2609	22077	18215
OME	ONT	ORD	ORF	OTZ	PBI	PDX	PFN	PHF	PHL	PHX
1013	40062	293422	18198	1022	22558	48049	1320	1016	91577	157155
PIA	PIT	PNS	PSC	PSG	PSP	PVD	PWM	RAP	RDU	RIC
76	126190	6621	1454	725	3132	12530	5451	1077	27642	15990
RNO	ROA	ROC	RST	RSW	SAN	SAT	SAV	SBA	SBN	SCC
20727	3312	16130	3514	14990	65731	38060	5763	1480	3560	632
SDF	SEA	SFO	SGF	SHV	SIT	SJC	SJU	SLC	SMF	SNA
24094	92016	129137	1575	4552	1311	39283	21458	84056	36051	36159
SPN	SRQ	STL	STT	STX	SUX	SWF	SYR	TLH	TOL	TPA
105	6863	174847	2874	1059	710	4402	12592	4668	2461	45782
TRI	TUL	TUS	TVC	TWF	TYS	VPS	WRG	YAK		
3056	20795	18549	1055	23	9509	1094	726	726		

In [16]:

```
%%R
sort(table(myDF$Origin))
```

project 2

LFT	LWB	TWF	ACK	MTJ	PIA	GST	DRO	SPN	ANI	KSM
2	4	23	40	45	76	98	102	105	133	133
GUC	FNT	DLG	OAJ	DUT	BFL	KOA	AKN	ADQ	HDN	HTS
243	248	367	369	379	455	455	475	504	556	615
SCC	BRO	EGE	EVV	SUX	PSG	CDV	WRG	YAK	BQN	CSG
632	687	690	709	710	725	726	726	726	734	971
BET	OME	PHF	FCA	OTZ	BRW	TVC	ITH	STX	RAP	HVN
981	1013	1016	1018	1022	1024	1055	1056	1059	1077	1081
MOT	HLN	VPS	MRY	IDA	SIT	PFN	MFR	ELM	LSE	LGB
1091	1094	1094	1205	1298	1311	1320	1342	1392	1392	1401
GFK	BIS	BGM	JAC	ERI	PSC	GNV	SBA	CAK	SGF	BGR
1414	1417	1422	1424	1454	1454	1473	1480	1574	1575	1589
MLI	DLH	MSO	CRW	EUG	FAY	ILM	BZN	MLU	FAR	LAN
1641	1813	1822	2031	2173	2188	2204	2205	2209	2339	2365
AGS	GTF	TOL	AZO	MGM	FAT	KTN	AVP	OGG	MYR	GRB
2387	2428	2461	2490	2498	2507	2589	2605	2609	2720	2843
STT	AVL	TRI	ISP	PSP	MLB	ROA	FWA	BTV	CHA	LNK
2874	2877	3056	3111	3132	3171	3312	3354	3362	3387	3488
RST	SBN	BIL	MBS	MFE	JNU	CRP	AMA	FSD	DAB	SWF
3514	3560	3701	3861	4057	4130	4225	4294	4343	4364	4402
SHV	MHT	MSN	TLH	BTR	FAI	JAN	LEX	CID	MOB	PWM
4552	4629	4640	4668	4741	4843	4945	5070	5318	5378	5451
CAE	HRL	SAV	HPN	LBB	PNS	SRQ	HSV	ICT	ABE	MAF
5714	5752	5763	6228	6338	6621	6863	6957	7051	7198	7512
MDT	GSP	GRR	DSM	CHS	TYS	BOI	ALB	COS	GEG	PVD
8648	8835	9171	9177	9228	9509	10174	11997	12077	12230	12530
HNL	SYR	DAY	RSW	RIC	ROC	LIT	MKE	BHM	JAX	ORF
12537	12592	14524	14990	15990	16130	16652	17172	18118	18184	18198
OMA	TUS	ANC	BUF	RNO	TUL	SJU	OKC	PBI	SDF	BDL
18215	18549	19040	19932	20727	20795	21458	22077	22558	24094	25771
RDU	ELP	BUR	GSO	IAD	FLL	SMF	SNA	AUS	ABQ	IND
27642	28728	31472	32676	33053	33641	36051	36159	36474	36871	37092
SAT	CMH	SJC	ONT	BNA	TPA	JFK	DAL	MDW	PDX	MSY
38060	38271	39283	40062	43382	45782	46187	46570	47985	48049	49179
MEM	MCI	HOU	OAK	SAN	BWI	CLE	MCO	MIA	CVG	SLC
53388	56727	57492	65541	65731	66154	66888	67900	70581	76629	84056
DCA	PHL	SEA	BOS	LGA	EWR	LAS	IAH	PIT	SFO	DEN
86266	91577	92016	92852	99225	109592	112181	121029	126190	129137	130327
CLT	MSP	DTW	PHX	STL	LAX	ATL	DFW	ORD		
131417	136436	147610	157155	174847	179909	230470	268127	293422		

In [17]:

```
%%R
table(myDF$Dest)
```

	ABE	ABQ	ACK	ADQ	AGS	AKN	ALB	AMA	ANC	ANI	ATL
7192	36865	39	504	2386	475	11992	4289	19047	133	230485	
	ATW	AUS	AVL	AVP	AZO	BDL	BET	BFL	BGM	BGR	BHM
1	36462	2877	2608	2489	25766	981	454	1422	1592	18121	
BIL	BIS	BNA	BOI	BOS	BQN	BRO	BRW	BTR	BTV	BUF	
3700	1418	43374	10176	92738	733	688	1024	4737	3362	19931	
BUR	BWI	BZN	CAE	CAK	CDV	CHA	CHS	CID	CLE	CLT	
31471	66443	2204	5712	1574	726	3389	9226	5321	66874	131420	
CMH	COS	CRP	CRW	CSG	CVG	DAB	DAL	DAY	DCA	DEN	
38283	12081	4223	2030	971	76605	4363	46570	14524	86239	130322	
DFW	DLG	DLH	DRO	DSM	DTW	DUT	EGE	ELM	ELP	ERI	
268159	367	1813	102	9174	147570	379	689	1392	28724	1454	
EUG	EVV	EWR	FAI	FAR	FAT	FAY	FCA	FLL	FNT	FSD	
2173	708	109570	4844	2340	2503	2188	1019	33632	247	4345	
FWA	GEG	GFK	GNV	GRB	GRR	GSO	GSP	GST	GTF	GUC	
3355	12234	1415	1475	2844	9168	32671	8833	98	2430	241	
GUM	HDN	HLN	HNL	HOU	HPN	HRL	HSV	HTS	HVN	IAD	
105	557	1094	12539	57483	6228	5752	6957	614	1081	33053	
IAH	ICT	IDA	ILM	IND	ISP	ITH	JAC	JAN	JAX	JFK	
121026	7050	1298	2204	37103	3110	1056	1421	4946	18183	46194	
JNU	KOA	KSM	KTN	LAN	LAS	LAX	LBB	LEX	LGA	LGB	
4130	456	133	2589	2366	112200	179950	6339	5071	99229	1401	
LIT	LNK	LSE	LWB	MAF	MBS	MCI	MCO	MDT	MDW	MEM	
16651	3486	1393	4	7513	3860	56733	67915	8652	47976	53402	
MFE	MFR	MGM	MHT	MIA	MKE	MLB	MLI	MLU	MOB	MOT	
4060	1341	2496	4628	70609	17176	3173	1641	2209	5373	1092	
MRY	MSN	MSO	MSP	MSY	MTJ	MYR	OAJ	OAK	OGG	OKC	
1206	4642	1822	136419	49170	45	2721	368	65538	2609	22064	
OMA	OME	ONT	ORD	ORF	OTZ	PBI	PDX	PFN	PHF	PHL	
18217	1013	40057	293472	18200	1022	22543	48042	1321	1014	91700	
PHX	PIA	PIT	PNS	PSC	PSG	PSP	PVD	PWM	RAP	RDU	
157158	75	126225	6621	1453	726	3133	12536	5452	1077	27646	
RIC	RNO	ROA	ROC	RST	RSW	SAN	SAT	SAV	SBA	SBN	
15991	20723	3312	16105	3515	14997	65724	38056	5762	1479	3562	
SCC	SDF	SEA	SFO	SGF	SHV	SIT	SJC	SJU	SLC	SMF	
632	24098	92016	129158	1573	4546	1310	39276	21082	84064	36056	
SNA	SRQ	STL	STT	STX	SUX	SWF	SYR	TLH	TOL	TPA	
36158	6860	174846	2946	1059	710	4401	12590	4667	2461	45760	
TRI	TUL	TUS	TVC	TWF	TYS	VPS	WRG	YAK			
3057	20805	18546	1056	23	9512	1094	725	726			

```
In [7]: %%R
table(myDF$Dest)[ "IND" ]
```

```
IND
37103
```

```
In [18]: %%R
sort(table(myDF$Dest))
```

project 2

ATW	LWB	TWF	ACK	MTJ	PIA	GST	DRO	GUM	ANI	KSM
1	4	23	39	45	75	98	102	105	133	133
GUC	FNT	DLG	OAJ	DUT	BFL	KOA	AKN	ADQ	HDN	HTS
241	247	367	368	379	454	456	475	504	557	614
SCC	BRO	EGE	EVV	SUX	WRG	CDV	PSG	YAK	BQN	CSG
632	688	689	708	710	725	726	726	726	733	971
BET	OME	PHF	FCA	OTZ	BRW	ITH	TVC	STX	RAP	HVN
981	1013	1014	1019	1022	1024	1056	1056	1059	1077	1081
MOT	HLN	VPS	MRY	IDA	SIT	PFN	MFR	ELM	LSE	LGB
1092	1094	1094	1206	1298	1310	1321	1341	1392	1393	1401
GFK	BIS	JAC	BGM	PSC	ERI	GNV	SBA	SGF	CAK	BGR
1415	1418	1421	1422	1453	1454	1475	1479	1573	1574	1592
MLI	DLH	MSO	CRW	EUG	FAY	BZN	ILM	MLU	FAR	LAN
1641	1813	1822	2030	2173	2188	2204	2204	2209	2340	2366
AGS	GTF	TOL	AZO	MGM	FAT	KTN	AVP	OGG	MYR	GRB
2386	2430	2461	2489	2496	2503	2589	2608	2609	2721	2844
AVL	STT	TRI	ISP	PSP	MLB	ROA	FWA	BTV	CHA	LNK
2877	2946	3057	3110	3133	3173	3312	3355	3362	3389	3486
RST	SBN	BIL	MBS	MFE	JNU	CRP	AMA	FSD	DAB	SWF
3515	3562	3700	3860	4060	4130	4223	4289	4345	4363	4401
SHV	MHT	MSN	TLH	BTR	FAI	JAN	LEX	CID	MOB	PWM
4546	4628	4642	4667	4737	4844	4946	5071	5321	5373	5452
CAE	HRL	SAV	HPN	LBB	PNS	SRQ	HSV	ICT	ABE	MAF
5712	5752	5762	6228	6339	6621	6860	6957	7050	7192	7513
MDT	GSP	GRR	DSM	CHS	TYS	BOI	ALB	COS	GEG	PVD
8652	8833	9168	9174	9226	9512	10176	11992	12081	12234	12536
HNL	SYR	DAY	RSW	RIC	ROC	LIT	MKE	BHM	JAX	ORF
12539	12590	14524	14997	15991	16105	16651	17176	18121	18183	18200
OMA	TUS	ANC	BUF	RNO	TUL	SJU	OKC	PBI	SDF	BDL
18217	18546	19047	19931	20723	20805	21082	22064	22543	24098	25766
RDU	ELP	BUR	GSO	IAD	FLL	SMF	SNA	AUS	ABQ	IND
27646	28724	31471	32671	33053	33632	36056	36158	36462	36865	37103
SAT	CMH	SJC	ONT	BNA	TPA	JFK	DAL	MDW	PDX	MSY
38056	38283	39276	40057	43374	45760	46194	46570	47976	48042	49170
MEM	MCI	HOU	OAK	SAN	BWI	CLE	MCO	MIA	CVG	SLC
53402	56733	57483	65538	65724	66443	66874	67915	70609	76605	84064
DCA	PHL	SEA	BOS	LGA	EWR	LAS	IAH	PIT	SFO	DEN
86239	91700	92016	92738	99229	109570	112200	121026	126225	129158	130322
CLT	MSP	DTW	PHX	STL	LAX	ATL	DFW	ORD		
131420	136419	147570	157158	174846	179950	230485	268159	293472		

In [8]:

```
%%R
table(myDF$Dest)[ "IND"]
```

```
IND
37103
```

In [19]:

```
%%R
head(myDF$Origin, row=894)
```

```
[1] ORD ORD ORD ORD ORD ORD
218 Levels: ABE ABQ ACK ADQ AGS AKN ALB AMA ANC ANI ATL AUS AVL AVP AZO ... YAK
```

In [6]:

```
%%R
head(myDF, row=894)
```

	Year	Month	DayofMonth	DayOfWeek	DepTime	CRSDepTime	ArrTime	CRSArrTime
1	1995	1	6	5	657	645	952	937
2	1995	1	7	6	648	645	938	937
3	1995	1	8	7	649	645	932	937
4	1995	1	9	1	645	645	928	937
5	1995	1	10	2	645	645	931	937
6	1995	1	11	3	646	645	929	937
	UniqueCarrier	FlightNum	TailNum	ActualElapsedTime	CRSElapsedTime	AirTime		
1	UA	482	N7298U		115		112	83
2	UA	482	N7449U		110		112	88
3	UA	482	N7453U		103		112	83
4	UA	482	N7288U		103		112	84
5	UA	482	N7275U		106		112	82
6	UA	482	N7281U		103		112	85
	ArrDelay	DepDelay	Origin	Dest	Distance	TaxiIn	TaxiOut	Cancelled
1	15	12	ORD	PHL	678	7	25	0
2	1	3	ORD	PHL	678	5	17	0
3	-5	4	ORD	PHL	678	3	17	0
4	-9	0	ORD	PHL	678	3	16	0
5	-6	0	ORD	PHL	678	6	18	0
6	-8	1	ORD	PHL	678	5	13	0
	CancellationCode	Diverted	CarrierDelay	WeatherDelay	NASDelay	SecurityDelay		
1	NA	0	NA	NA	NA	NA	NA	NA
2	NA	0	NA	NA	NA	NA	NA	NA
3	NA	0	NA	NA	NA	NA	NA	NA
4	NA	0	NA	NA	NA	NA	NA	NA
5	NA	0	NA	NA	NA	NA	NA	NA
6	NA	0	NA	NA	NA	NA	NA	NA
	LateAircraftDelay							
1		NA						
2		NA						
3		NA						
4		NA						
5		NA						
6		NA						

In [23]:

```
%%R
head(myDF$Dest, row=894)
```

```
[1] PHL PHL PHL PHL PHL PHL
218 Levels: ABE ABQ ACK ADQ AGS AKN ALB AMA ANC ANI ATL ATW AUS AVL AVP ... YAK
```

In [7]:

```
%%R
head(myDF, row=894)
```

	Year	Month	DayofMonth	DayOfWeek	DepTime	CRSDepTime	ArrTime	CRSArrTime
1	1995	1	6	5	657	645	952	937
2	1995	1	7	6	648	645	938	937
3	1995	1	8	7	649	645	932	937
4	1995	1	9	1	645	645	928	937
5	1995	1	10	2	645	645	931	937
6	1995	1	11	3	646	645	929	937
	UniqueCarrier	FlightNum	TailNum	ActualElapsedTime	CRSElapsedTime	AirTime		
1	UA	482	N7298U		115		112	83
2	UA	482	N7449U		110		112	88
3	UA	482	N7453U		103		112	83
4	UA	482	N7288U		103		112	84
5	UA	482	N7275U		106		112	82
6	UA	482	N7281U		103		112	85
	ArrDelay	DepDelay	Origin	Dest	Distance	TaxiIn	TaxiOut	Cancelled
1	15	12	ORD	PHL	678	7	25	0
2	1	3	ORD	PHL	678	5	17	0
3	-5	4	ORD	PHL	678	3	17	0
4	-9	0	ORD	PHL	678	3	16	0
5	-6	0	ORD	PHL	678	6	18	0
6	-8	1	ORD	PHL	678	5	13	0
	CancellationCode	Diverted	CarrierDelay	WeatherDelay	NASDelay	SecurityDelay		
1	NA	0	NA	NA	NA	NA	NA	NA
2	NA	0	NA	NA	NA	NA	NA	NA
3	NA	0	NA	NA	NA	NA	NA	NA
4	NA	0	NA	NA	NA	NA	NA	NA
5	NA	0	NA	NA	NA	NA	NA	NA
6	NA	0	NA	NA	NA	NA	NA	NA
	LateAircraftDelay							
1		NA						
2		NA						
3		NA						
4		NA						
5		NA						
6		NA						

In [2]: %%R
head(myDF)

	Year	Month	DayofMonth	DayOfWeek	DepTime	CRSDepTime	ArrTime	CRSArrTime
1	1995	1	6	5	657	645	952	937
2	1995	1	7	6	648	645	938	937
3	1995	1	8	7	649	645	932	937
4	1995	1	9	1	645	645	928	937
5	1995	1	10	2	645	645	931	937
6	1995	1	11	3	646	645	929	937
	UniqueCarrier	FlightNum	TailNum	ActualElapsedTime	CRSElapsedTime	AirTime		
1	UA	482	N7298U		115	112	83	
2	UA	482	N7449U		110	112	88	
3	UA	482	N7453U		103	112	83	
4	UA	482	N7288U		103	112	84	
5	UA	482	N7275U		106	112	82	
6	UA	482	N7281U		103	112	85	
	ArrDelay	DepDelay	Origin	Dest	Distance	TaxiIn	TaxiOut	Cancelled
1	15	12	ORD	PHL	678	7	25	0
2	1	3	ORD	PHL	678	5	17	0
3	-5	4	ORD	PHL	678	3	17	0
4	-9	0	ORD	PHL	678	3	16	0
5	-6	0	ORD	PHL	678	6	18	0
6	-8	1	ORD	PHL	678	5	13	0
	CancellationCode	Diverted	CarrierDelay	WeatherDelay	NASDelay	SecurityDelay		
1	NA	0	NA	NA	NA	NA	NA	
2	NA	0	NA	NA	NA	NA	NA	
3	NA	0	NA	NA	NA	NA	NA	
4	NA	0	NA	NA	NA	NA	NA	
5	NA	0	NA	NA	NA	NA	NA	
6	NA	0	NA	NA	NA	NA	NA	
	LateAircraftDelay							
1	NA							
2	NA							
3	NA							
4	NA							
5	NA							
6	NA							

In [3]:

```
%%R
head(myDF$Dist)
```

```
[1] 678 678 678 678 678 678
```

In [4]:

```
%%R
tail(myDF$Dist)
```

```
[1] 770 770 770 770 770 770
```

In [5]:

```
%%R
summary(myDF$Dist)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
11.0	316.0	545.0	695.5	934.0	4502.0	5987

In [19]:

```
%%R
length(myDF$Dist[myDF$Dist < 200])
# this method includes the NA values ( it keeps the NA values around, incase you want
```

```
[1] 475141
```

In [16]:

```
%%R
head(myDF$Dist < 200)
```

```
[1] FALSE FALSE FALSE FALSE FALSE FALSE
```

In [17]:

```
%%R
head(myDF$Dist >= 200)
```

```
[1] TRUE TRUE TRUE TRUE TRUE TRUE
```

In [20]:

```
%%R
sum(myDF$Dist < 200)
# when i tried adding TRUEand FALSE values
# the TRUE values become 1
# the FALSE value become 0
# so the sum is the number of true values
```

```
[1] NA
```

In [22]:

```
%%R
sum(myDF$Dist < 200, na.rm=TRUE)
# when i tried adding TRUEand FALSE values
# the TRUE values become 1
# the FALSE value become 0
# so the sum is the number of true values
```

```
[1] 469154
```

In [23]:

```
%%R
469154 + 5987
```

```
[1] 475141
```

In [25]:

```
%%R
sum(myDF$Dist < 200, na.rm=TRUE)
```

```
[1] 469154
```

In [26]:

```
%%R
sum(myDF$Dist >= 200, na.rm=TRUE)
```

```
[1] 4852294
```

In [28]:

```
%%R
sum(is.na(myDF$Dist))
```

```
[1] 5987
```

In [29]:

```
%%R
469154 + 4852294 + 5987
```

```
[1] 5327435
```

In []:

How many flights departed from Indianapolis (IND) in 1995? 37092 How many flights landed there? 37103 Consider the flight data from row 894 the data frame. What airport did it depart from? ORD Where did it arrive? PHL How many flights have a distance of less than 200 miles? 475141

Question 4

In [3]:

```
%%R  
table(myDF$UniqueCarrier)
```

AA	AS	CO	DL	HP	NW	TW	UA	US	WN
688471	135124	434425	884019	194508	521693	272452	724807	778835	693101

In [4]:

```
%%R  
sort(table(myDF$UniqueCarrier))
```

AS	HP	TW	CO	NW	AA	WN	UA	US	DL
135124	194508	272452	434425	521693	688471	693101	724807	778835	884019

In [5]:

```
%%R  
sort(table(myDF$UniqueCarrier), decreasing=TRUE)
```

DL	US	UA	WN	AA	NW	CO	TW	HP	AS
884019	778835	724807	693101	688471	521693	434425	272452	194508	135124

In [6]:

```
%%R  
head(sort(table(myDF$UniqueCarrier), decreasing=TRUE), n=5)
```

DL	US	UA	WN	AA
884019	778835	724807	693101	688471

In [7]:

```
%%R  
head(sort(table(myDF$UniqueCarrier), decreasing=TRUE), n=3)
```

DL	US	UA
884019	778835	724807

In [8]:

```
%%R  
head(myDF)
```

	Year	Month	DayofMonth	DayOfWeek	DepTime	CRSDepTime	ArrTime	CRSArrTime
1	1995	1	6	5	657	645	952	937
2	1995	1	7	6	648	645	938	937
3	1995	1	8	7	649	645	932	937
4	1995	1	9	1	645	645	928	937
5	1995	1	10	2	645	645	931	937
6	1995	1	11	3	646	645	929	937
	UniqueCarrier	FlightNum	TailNum	ActualElapsedTime	CRSElapsedTime	AirTime		
1	UA	482	N7298U		115		112	83
2	UA	482	N7449U		110		112	88
3	UA	482	N7453U		103		112	83
4	UA	482	N7288U		103		112	84
5	UA	482	N7275U		106		112	82
6	UA	482	N7281U		103		112	85
	ArrDelay	DepDelay	Origin	Dest	Distance	TaxiIn	TaxiOut	Cancelled
1	15	12	ORD	PHL	678	7	25	0
2	1	3	ORD	PHL	678	5	17	0
3	-5	4	ORD	PHL	678	3	17	0
4	-9	0	ORD	PHL	678	3	16	0
5	-6	0	ORD	PHL	678	6	18	0
6	-8	1	ORD	PHL	678	5	13	0
	CancellationCode	Diverted	CarrierDelay	WeatherDelay	NASDelay	SecurityDelay		
1	NA	0	NA	NA	NA	NA	NA	NA
2	NA	0	NA	NA	NA	NA	NA	NA
3	NA	0	NA	NA	NA	NA	NA	NA
4	NA	0	NA	NA	NA	NA	NA	NA
5	NA	0	NA	NA	NA	NA	NA	NA
6	NA	0	NA	NA	NA	NA	NA	NA
	LateAircraftDelay							
1		NA						
2		NA						
3		NA						
4		NA						
5		NA						
6		NA						

In [9]:

```
%%R  
head(myDF$TailNum)
```

```
[1] N7298U N7449U N7453U N7288U N7275U N7281U  
3508 Levels: 112 EI-BOM EI-BON EI-BWD EI-CAL EI-CAM EI-CIW N000AA ... UNKNOW
```

In [14]:

```
%%R  
head(sort(table(myDF$TailNum),decreasing=TRUE))
```

UNKNOW	N67	N89	N83	N71	N103
65072	3485	3479	3476	3474	3463

In [15]:

```
%%R  
head(sort(table(myDF$TailNum),decreasing=TRUE), n=50)
```

UNKNOW	N67	N89	N83	N71	N103	N68	N102	N91	N59	N64
65072	3485	3479	3476	3474	3463	3458	3448	3448	3447	3434
N92	N20	N23	N86	N130	N62	N61	N73	N82	N511	N57
3433	3430	3427	3426	3424	3424	3414	3409	3407	3401	3400
N104	N514	N93	N28	N74	N326	N519	N520	N85	N359	N90
3397	3396	3393	3387	3384	3383	3383	3383	3380	3372	3371
N522	N63	N95	N358	N105	N526	N54	N512	N504	N523	N361
3368	3365	3363	3360	3359	3359	3359	3353	3349	3346	3343
N503	N515	N94	N325	N323	N60					
3340	3339	3335	3332	3331	3331					

In []:

a. Rank the airline companies

AS	HP	TW	CO	NW	AA	WN	UA	US	DL
135124	194508	272452	434425	521693	688471	693101	724807	778835	884019

b. Which are the three most popular airlines from 1995?

DL	US	UA
884019	778835	724807

c. Now find the ten airplanes that had the most flights in 1995. List them in order, from most popular to least popular. Do you notice anything unusual about the results?

N64	N67	N89	N83	N71	N103	N68	N102	N91	N59
3434	3485	3479	3476	3474	3463	3458	3448	3448	3447

their differences or variation from each other is pretty minimal.

Question 5

In [2]:

```
%%R
table(myDF$Origin)
```

ABE	ABQ	ACK	ADQ	AGS	AKN	ALB	AMA	ANC	ANI	ATL
7198	36871	40	504	2387	475	11997	4294	19040	133	230470
AUS	AVL	AVP	AZO	BDL	BET	BFL	BGM	BGR	BHM	BIL
36474	2877	2605	2490	25771	981	455	1422	1589	18118	3701
BIS	BNA	BOI	BOS	BQN	BRO	BRW	BTR	BTW	BUF	BUR
1417	43382	10174	92852	734	687	1024	4741	3362	19932	31472
BWI	BZN	CAE	CAK	CDV	CHA	CHS	CID	CLE	CLT	CMH
66154	2205	5714	1574	726	3387	9228	5318	66888	131417	38271
COS	CRP	CRW	CSG	CVG	DAB	DAL	DAY	DCA	DEN	DFW
12077	4225	2031	971	76629	4364	46570	14524	86266	130327	268127
DLG	DLH	DRO	DSM	DTW	DUT	EGE	ELM	ELP	ERI	EUG
367	1813	102	9177	147610	379	690	1392	28728	1454	2173
EVV	EWR	FAI	FAR	FAT	FAY	FCA	FLL	FNT	FSD	FWA
709	109592	4843	2339	2507	2188	1018	33641	248	4343	3354
GEG	GFK	GNV	GRB	GRR	GSO	GSP	GST	GTF	GUC	HDN
12230	1414	1473	2843	9171	32676	8835	98	2428	243	556
HLN	HNL	HOU	HPN	HRL	HSV	HTS	HVN	IAD	IAH	ICT
1094	12537	57492	6228	5752	6957	615	1081	33053	121029	7051
IDA	ILM	IND	ISP	ITH	JAC	JAN	JAX	JFK	JNU	KOA
1298	2204	37092	3111	1056	1424	4945	18184	46187	4130	455
KSM	KTN	LAN	LAS	LAX	LBB	LEX	LFT	LGA	LGB	LIT
133	2589	2365	112181	179909	6338	5070	2	99225	1401	16652
LNK	LSE	LWB	MAF	MBS	MCI	MCO	MDT	MDW	MEM	MFE
3488	1392	4	7512	3861	56727	67900	8648	47985	53388	4057
MFR	MGM	MHT	MIA	MKE	MLB	MLI	MLU	MOB	MOT	MRY
1342	2498	4629	70581	17172	3171	1641	2209	5378	1091	1205
MSN	MSO	MSP	MSY	MTJ	MYR	OAJ	OAK	OGG	OKC	OMA
4640	1822	136436	49179	45	2720	369	65541	2609	22077	18215
OME	ONT	ORD	ORF	OTZ	PBI	PDX	PFN	PHF	PHL	PHX
1013	40062	293422	18198	1022	22558	48049	1320	1016	91577	157155
PIA	PIT	PNS	PSC	PSG	PSP	PVD	PWM	RAP	RDU	RIC
76	126190	6621	1454	725	3132	12530	5451	1077	27642	15990
RNO	ROA	ROC	RST	RSW	SAN	SAT	SAV	SBA	SBN	SCC
20727	3312	16130	3514	14990	65731	38060	5763	1480	3560	632
SDF	SEA	SFO	SGF	SHV	SIT	SJC	SJU	SLC	SMF	SNA
24094	92016	129137	1575	4552	1311	39283	21458	84056	36051	36159
SPN	SRQ	STL	STT	STX	SUX	SWF	SYR	TLH	TOL	TPA
105	6863	174847	2874	1059	710	4402	12592	4668	2461	45782
TRI	TUL	TUS	TVC	TWF	TYS	VPS	WRG	YAK		
3056	20795	18549	1055	23	9509	1094	726	726		

In [3]:

```
%%R
sort(table(myDF$Origin), decreasing=TRUE)
```

	ORD	DFW	ATL	LAX	STL	PHX	DTW	MSP	CLT	DEN	SFO
293422	268127	230470	179909	174847	157155	147610	136436	131417	130327	129137	
PIT	IAH	LAS	EWR	LGA	BOS	SEA	PHL	DCA	SLC	CVG	
126190	121029	112181	109592	99225	92852	92016	91577	86266	84056	76629	
MIA	MCO	CLE	BWI	SAN	OAK	HOU	MCI	MEM	MSY	PDX	
70581	67900	66888	66154	65731	65541	57492	56727	53388	49179	48049	
MDW	DAL	JFK	TPA	BNA	ONT	SJC	CMH	SAT	IND	ABQ	
47985	46570	46187	45782	43382	40062	39283	38271	38060	37092	36871	
AUS	SNA	SMF	FLL	IAD	GSO	BUR	ELP	RDU	BDL	SDF	
36474	36159	36051	33641	33053	32676	31472	28728	27642	25771	24094	
PBI	OKC	SJU	TUL	RNO	BUF	ANC	TUS	OMA	ORF	JAX	
22558	22077	21458	20795	20727	19932	19040	18549	18215	18198	18184	
BHM	MKE	LIT	ROC	RIC	RSW	DAY	SYR	HNL	PVD	GEG	
18118	17172	16652	16130	15990	14990	14524	12592	12537	12530	12230	
COS	ALB	BOI	TYC	CHS	DSM	GRR	GSP	MDT	MAF	ABE	
12077	11997	10174	9509	9228	9177	9171	8835	8648	7512	7198	
ICT	HSV	SRQ	PNS	LBB	HPN	SAV	HRL	CAE	PWM	MOB	
7051	6957	6863	6621	6338	6228	5763	5752	5714	5451	5378	
CID	LEX	JAN	FAI	BTR	TLH	MSN	MHT	SHV	SWF	DAB	
5318	5070	4945	4843	4741	4668	4640	4629	4552	4402	4364	
FSD	AMA	CRP	JNU	MFE	MBS	BIL	SBN	RST	LNK	CHA	
4343	4294	4225	4130	4057	3861	3701	3560	3514	3488	3387	
BTV	FWA	ROA	MLB	PSP	ISP	TRI	AVL	STT	GRB	MYR	
3362	3354	3312	3171	3132	3111	3056	2877	2874	2843	2720	
OGG	AVP	KTN	FAT	MGM	AZO	TOL	GTF	AGS	LAN	FAR	
2609	2605	2589	2507	2498	2490	2461	2428	2387	2365	2339	
MLU	BZN	ILM	FAY	EUG	CRW	MSO	DLH	MLI	BGR	SGF	
2209	2205	2204	2188	2173	2031	1822	1813	1641	1589	1575	
CAK	SBA	GNV	ERI	PSC	JAC	BGM	BIS	GFK	LGB	ELM	
1574	1480	1473	1454	1454	1424	1422	1417	1414	1401	1392	
LSE	MFR	PFN	SIT	IDA	MRY	HLN	VPS	MOT	HVN	RAP	
1392	1342	1320	1311	1298	1205	1094	1094	1091	1081	1077	
STX	ITH	TVC	BRW	OTZ	FCA	PHF	OME	BET	CSG	BQN	
1059	1056	1055	1024	1022	1018	1016	1013	981	971	734	
CDV	WRG	YAK	PSG	SUX	EVV	EGE	BRO	SCC	HTS	HDN	
726	726	726	725	710	709	690	687	632	615	556	
ADQ	AKN	BFL	KOA	DUT	OAJ	DLG	FNT	GUC	ANI	KSM	
504	475	455	455	379	369	367	248	243	133	133	
SPN	DRO	GST	PIA	MTJ	ACK	TWF	LWB	LFT			
105	102	98	76	45	40	23	4	2			

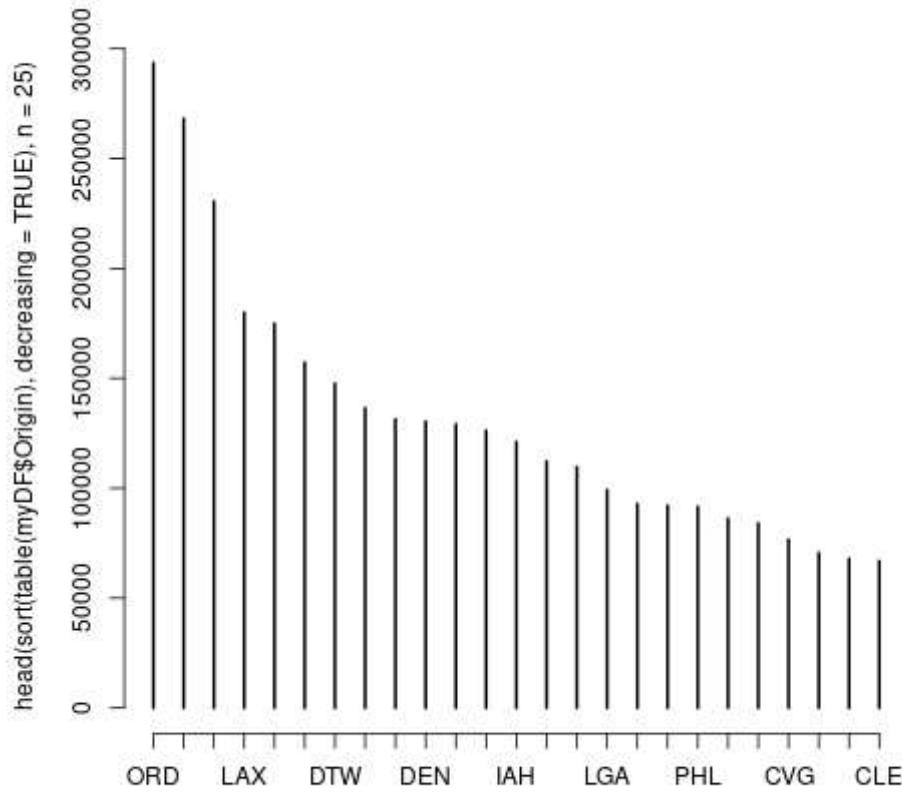
In [5]:

```
%%R
head(sort(table(myDF$Origin), decreasing=TRUE), n=25)
```

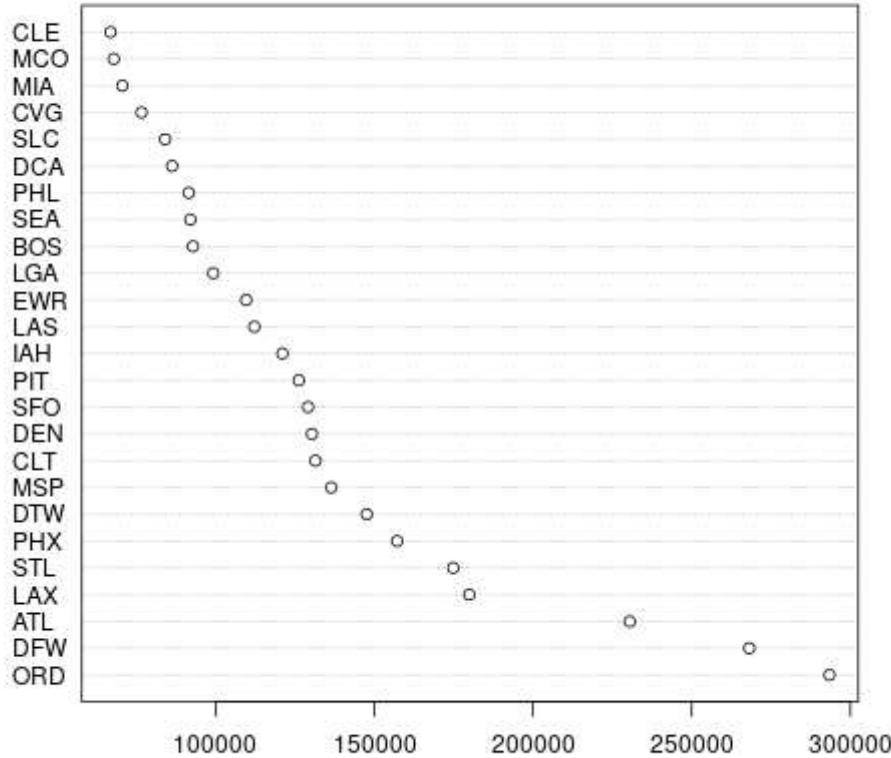
	ORD	DFW	ATL	LAX	STL	PHX	DTW	MSP	CLT	DEN	SFO
293422	268127	230470	179909	174847	157155	147610	136436	131417	130327	129137	
PIT	IAH	LAS	EWR	LGA	BOS	SEA	PHL	DCA	SLC	CVG	
126190	121029	112181	109592	99225	92852	92016	91577	86266	84056	76629	
MIA	MCO	CLE									
70581	67900	66888									

In [11]:

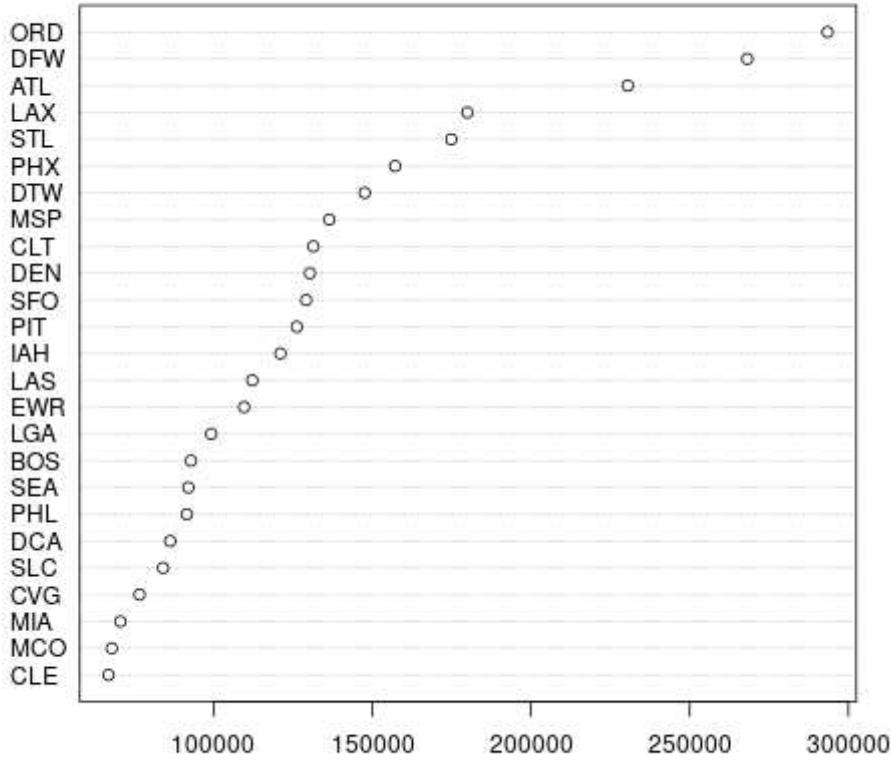
```
%%R
plot(head(sort(table(myDF$Origin), decreasing=TRUE), n=25))
```



```
In [12]: %%R  
dotchart(head(sort(table(myDF$Origin), decreasing=TRUE),n=25))
```



```
In [13]: %%R  
dotchart( rev(head(sort(table(myDF$Origin),decreasing=TRUE),n=25) ))
```



In []:

In []:

In []:

Describe any patterns you may see in the plot and the dotchart. It has a snake like shape.

Pledge

By submitting this work I hereby pledge that this is my own, personal work. I've acknowledged in the designated place at the top of this file all sources that I used to complete said work, including but not limited to: online resources, books, and electronic communications. I've noted all collaboration with fellow students and/or TA's. I did not copy or plagiarize another's work.

As a Boilermaker pursuing academic excellence, I pledge to be honest and true in all that I do. Accountable together – We are Purdue.