

Stats 506, F18, Problem Set 1

Chen Xie, *chenxie@umich.edu*

October 01, 2018

Question 1

Part A

i. `cat recs2015_public_v3.csv |awk -F, '{if($2~/3/) {print $2}}' |wc -l`

Result:

2010

There are 2010 rows for region 3 in the RECS 2015 data set.

ii. `cat recs2015_public_v3.csv | cut -d, -f 1,475-571>recs_15_sub.csv`

The compressed data set is shown in the file “recs_15_sub.csv”.

Part B

i. `for i in 1 2 3 4
do
cat recs2015_public_v3.csv |
awk -F, '{if($2~/\'$i\'/) {print $2}}' |wc -l
done`

Result:

794

1327

2010

1555

The observations for region 1, 2, 3, 4 are 794, 1327, 2010, 1555, respectively.

ii. `cat recs2015_public_v3.csv |
awk -F, 'NR==1{print$2 $3};NR > 1{print $2 $3|"sort -V"}' |
uniq>region_division.txt`

The unique combinations of values from RECIIONC and DIVISION are (“1”“1”, “1”“2”, “2”“3”, “2”“4”, “3”“5”, “3”“6”, “3”“7”, “4”“8”, “4”“9”, “4”“10”). The result is written in the file “region_division.txt”.

Question 2

Part a

Table 1: **Q2_part_a.** *The Airlines* that were responsible for at least 1% of the flights departing any of the three NYC airports between Jan 1st and Oct 1st, 2013. Rows are sorted by the percent of flights.

Airlines	Percent(%)
United Air Lines Inc.	17.372
JetBlue Airways	16.208
ExpressJet Airlines Inc.	16.133
Delta Air Lines Inc.	14.276
American Airlines Inc.	9.755
Envoy Air	7.891
US Airways Inc.	6.124
Endeavor Air Inc.	5.413
Southwest Airlines Co.	3.605
Virgin America	1.505
AirTran Airways Corporation	1.011

Part b

Table 2: Number and Percent with CI in 2013

Airlines	Number	Percent	lower ci(%)	upper ci(%)
Southwest Airlines Co.	10143	3.60	3.54	3.67
Delta Air Lines Inc.	40168	14.28	14.15	14.40
Virgin America	4235	1.51	1.46	1.55
US Airways Inc.	17232	6.12	6.04	6.21
JetBlue Airways	45605	16.21	16.07	16.34
American Airlines Inc.	27447	9.75	9.65	9.86
AirTran Airways Corporation	2845	1.01	0.97	1.05
United Air Lines Inc.	48880	17.37	17.23	17.51
Envoy Air	22202	7.89	7.79	7.99
ExpressJet Airlines Inc.	45395	16.13	16.00	16.27
Endeavor Air Inc.	15232	5.41	5.33	5.50

The Table 2 is the table of number and percents (%) with 95% CI of annual flights in the first 10 months of 2013.

Table 3: Number and Percent with CI in 2014

Airlines	Number	Percent	lower ci(%)	upper ci(%)
Southwest Airlines Co.	11902	4.70	4.62	4.78
Delta Air Lines Inc.	41683	16.45	16.31	16.60
Virgin America	4797	1.89	1.84	1.95
US Airways Inc.	16750	6.61	6.52	6.71
JetBlue Airways	44479	17.56	17.41	17.71
American Airlines Inc.	26302	10.38	10.26	10.50
AirTran Airways Corporation	1251	0.49	0.47	0.52
United Air Lines Inc.	46267	18.26	18.11	18.42
Envoy Air	18559	7.33	7.22	7.43
ExpressJet Airlines Inc.	39819	15.72	15.58	15.86
Endeavor Air Inc.	NA	NA	NA	NA

The Table 3 is the table of number and percents (%) with 95% CI of annual flights in the first 10 months 2014.

Table 4: Change in Number and Percent between 2013 and 2014

Airlines	Change in Number	Change in Percent(%)	lower ci(%)	upper ci(%)
Southwest Airlines Co.	1759	1.09	0.99	1.20
Delta Air Lines Inc.	1515	2.18	1.99	2.37
Virgin America	562	0.39	0.32	0.46
US Airways Inc.	-482	0.49	0.36	0.62
JetBlue Airways	-1126	1.35	1.15	1.55
American Airlines Inc.	-1145	0.63	0.47	0.79
AirTran Airways Corporation	-1594	-0.52	-0.56	-0.47
United Air Lines Inc.	-2613	0.89	0.69	1.10
Envoy Air	-3643	-0.56	-0.71	-0.42
ExpressJet Airlines Inc.	-5576	-0.41	-0.61	-0.22
Endeavor Air Inc.	NA	NA	NA	NA

The Table 4 is the table of changes in numbers between 2013 and 2014, as well as change in percent (%) with 95% CI.

It seems that Southwest Airlines has the largest increase in number, and ExpressJet Airlines has the largest decrease in number. And Delta Airlines shows the largest increase in percent, while Envoy Air shows the largest decrease in percent. The values for Endeavo Air for 2014 is missing.

Some airlines show an increase in the percent of flights but a decrease in the number, like United Airlines, American Airlines, US Airways, and JetBule Airways. As the total number of flights decrease more than the number decrease of these airlines.

Part c

Table 5: Percent of Flights Part 1

origin	9E	AA	AS	B6	DL	EV	F9	FL
EWB	1.05	2.89	0.59	5.43	3.59	36.36	NA	NA
JFK	13.17	12.39	NA	37.81	18.60	1.27	NA	NA
LGA	2.43	14.77	NA	5.73	22.04	8.43	0.65	3.11

Table 6: Percent of Flights Part 2

origin	HA	MQ	OO	UA	US	VX	WN	YV
EWB	NA	1.88	0.00	38.14	3.65	1.30	5.12	NA
JFK	0.31	6.46	NA	4.07	2.69	3.23	NA	NA
LGA	NA	16.17	0.02	7.69	12.55	NA	5.82	0.57

The Table 5 and 6 are the tables showing the percent (%) of flights each airlines is responsible for each of the airports in 2013. For showing precisely, I use abbreviation for airports names and airlines names.

Table 7: Percent of Flights with CI

Airports	Airlines	Percent(%)	lower ci(%)	upper ci(%)
Newark Liberty Intl	Endeavor Air Inc.	1.05	0.99	1.11
Newark Liberty Intl	American Airlines Inc.	2.89	2.79	2.98
Newark Liberty Intl	JetBlue Airways	5.43	5.30	5.55
Newark Liberty Intl	Delta Air Lines Inc.	3.59	3.49	3.70
Newark Liberty Intl	ExpressJet Airlines Inc.	36.36	36.09	36.63
Newark Liberty Intl	Envoy Air	1.88	1.81	1.96
Newark Liberty Intl	United Air Lines Inc.	38.14	37.87	38.41
Newark Liberty Intl	US Airways Inc.	3.65	3.54	3.75
Newark Liberty Intl	Virgin America	1.30	1.23	1.36
Newark Liberty Intl	Southwest Airlines Co.	5.12	5.00	5.25
John F Kennedy Intl	Endeavor Air Inc.	13.17	12.97	13.36
John F Kennedy Intl	American Airlines Inc.	12.39	12.19	12.58
John F Kennedy Intl	JetBlue Airways	37.81	37.53	38.10
John F Kennedy Intl	Delta Air Lines Inc.	18.60	18.37	18.83
John F Kennedy Intl	ExpressJet Airlines Inc.	1.27	1.20	1.33
John F Kennedy Intl	Envoy Air	6.46	6.32	6.61
John F Kennedy Intl	United Air Lines Inc.	4.07	3.96	4.19
John F Kennedy Intl	US Airways Inc.	2.69	2.60	2.79
John F Kennedy Intl	Virgin America	3.23	3.13	3.34
La Guardia	Endeavor Air Inc.	2.43	2.33	2.52
La Guardia	American Airlines Inc.	14.77	14.56	14.99
La Guardia	JetBlue Airways	5.73	5.59	5.88
La Guardia	Delta Air Lines Inc.	22.04	21.79	22.29
La Guardia	ExpressJet Airlines Inc.	8.43	8.26	8.60
La Guardia	AirTran Airways Corporation	3.11	3.01	3.22
La Guardia	Envoy Air	16.17	15.95	16.40
La Guardia	United Air Lines Inc.	7.69	7.52	7.85
La Guardia	US Airways Inc.	12.55	12.35	12.75
La Guardia	Southwest Airlines Co.	5.82	5.67	5.96

The Table 7 the table limited to the airlines in part a, showing that the percent (%) with 95% CI of flights each airline for each NYC airport.

Table 8: Largest Carrier

Airports	Airlines	Percent(%)	lower ci(%)	upper ci(%)
Newark Liberty Intl	United Air Lines Inc.	38.14	37.87	38.41
John F Kennedy Intl	JetBlue Airways	37.81	37.53	38.10
La Guardia	Delta Air Lines Inc.	22.04	21.79	22.29

The Table 8 is the table showing the largest carries at each NYC airport. For the airport Newark Liberty Intl, United Aialines is the largest carrier, while for John F kennedy Intl, the largest is JetBlue Airways, and for La Guardia, it is Delta Airlines.

Question 3

Part a

The Table 9 is the table showing estimated percent and BRR se of homes have stucco as the major wall material for each division. The division Mountain South has the highest proportion, and the division East South Central has the lowest proportion.

Table 9: Proportion of Stucco for each Divison with BRRse

Divison	Proportion of stucco(%)	BRR se
Mountain South	64.2479	4.4901
Pacific	44.5861	1.6874
Mountain North	16.6055	3.2432
South Atlantic	10.6197	1.4209
West North Central	4.8656	2.0277
West South Central	2.9873	0.6941
Middle Atlantic	2.0616	0.7344
New England	1.2265	0.8010
East North Central	0.6565	0.2834
East South Central	0.4230	0.4090

Part b

The Table 10 is the table showing estimated average and BRR se of total electricity usage in kilowatt in each division.

Table 10: Average Electricity Usage for each Divison with BRRse

Divison	Average eletricity usage (kw)	BRR se
East North Central	9128.66	203.63
East South Central	14535.97	620.63
Middle Atlantic	8465.44	201.10
Mountain North	8384.47	644.43
Mountain South	10442.02	1271.40
New England	7514.56	532.02
Pacific	8100.41	178.60
South Atlantic	13446.62	276.88
West North Central	10523.83	453.62
West South Central	14324.26	423.04

The Table 11 is the table showing estimated average and BRR se of total electricity usage in kilowatt for each urban type.

Table 11: Average Electricity Usage for Urban and Rural with BRRse

Urbantype	Average eletricity usage (kw)	BRR se
Rural	14164.87	390.05
Urban	9867.52	130.23

Part c

The Table 12 is the table showing estimates and BRR se of disparity between urban and rural areas of the proportion of homes with internet access for each division. The Division Mountain South has the largest disparity in proportion.

Table 12: Disparity of Proportion in Internet access between Urban and Rural for each Division with BRRse

Division	Disparity of the proportion(%)	BRR se
Mountain South	18.52	5.78
East South Central	9.33	5.47
West North Central	7.68	5.17
Mountain North	5.50	5.96
West South Central	5.10	3.77
Pacific	3.43	3.80
South Atlantic	3.26	3.47
New England	1.78	1.72
East North Central	0.04	2.67
Middle Atlantic	-1.95	3.25