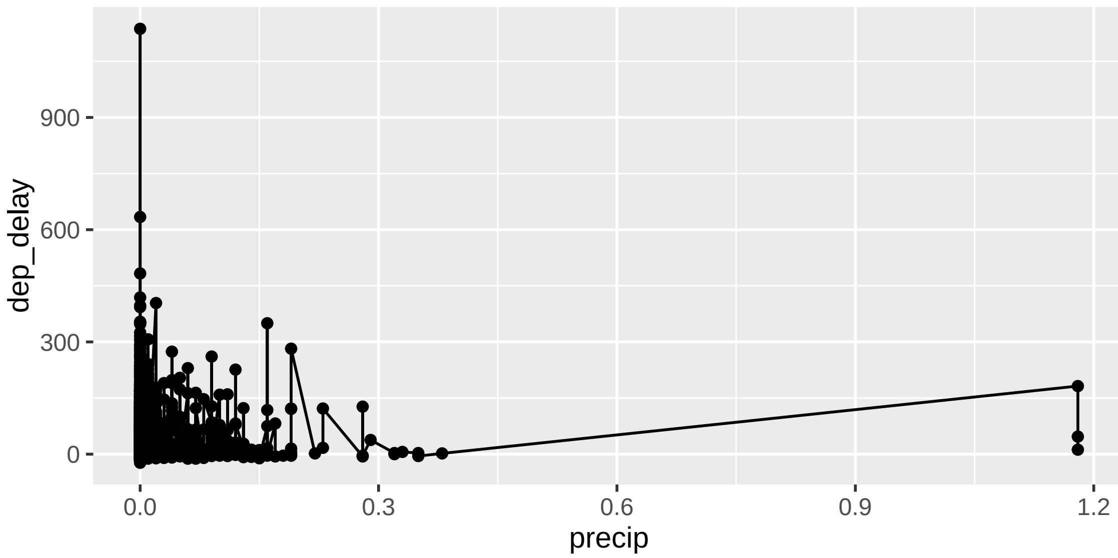
Econ 294A Final

Richard Ronson

March 14, 2016

Part A



The plot above illustrates the most significant relationship of part A. This is between precipitation and depature delay.

The regression results of Part A aimed to figure out the relationship between various

weather conditions with both delayed and cancelled flights.

All of the weather condtions were statistically significant at the 99% level in

determining delayed depatures. Likewise only a few weather conditons played no

signifcant role in leading to a cancelled flight. These conditions include: humidity,

wind speed and precipitation. I would like to take note of the most significant statistic on table one which was the effect of precipitation on depature delay. This

was that on average holding all else constant as precipitation increases by 1 unit

depature delay will increase by 49.29 mins. This far and away was the greatest find of

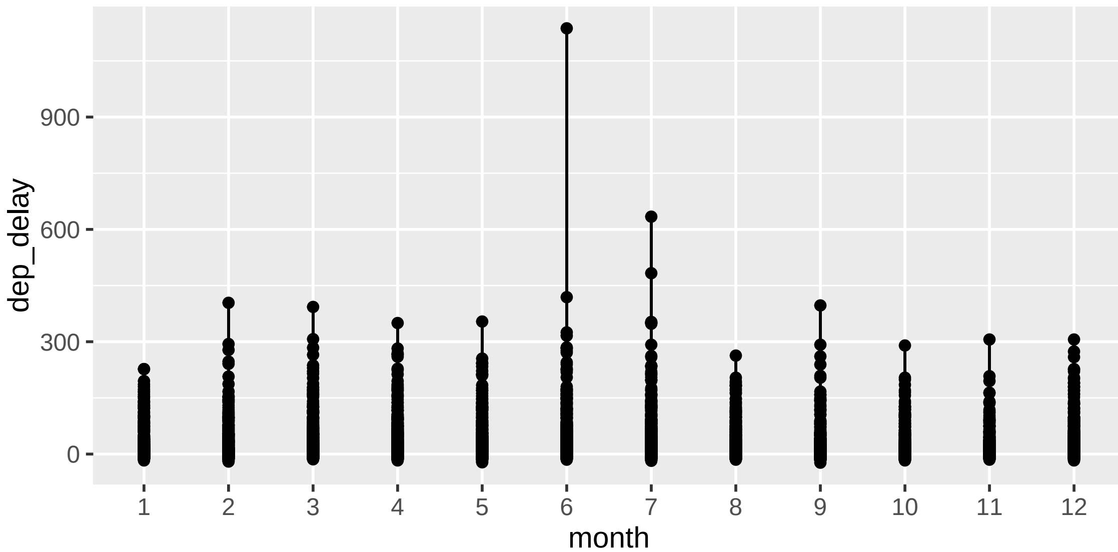
the regression results of part A both in statistical signifiance and magnitude.

(reference to table 1)

Table 1:

	Dependent variable:	
	dep_delay	canceled
	(1)	(2)
temp	0.189***	0.00004***
	(0.004)	(0.00000)
humid	−0.145***	−0.00000
	(0.004)	(0.00000)
wind_speed	0.028***	0.00000
	(0.004)	(0.00000)
precip	49.282***	−0.003
	(5.016)	(0.005)
pressure	−0.424***	−0.00005***
	(0.011)	(0.00001)
visib	−2.089***	−0.0002***
	(0.055)	(0.0001)
wind_dir	−0.005***	−0.00000***
	(0.001)	(0.00000)
Constant	461.473***	0.050***
	(10.911)	(0.010)
Observations	281,563	281,563
R^2	0.026	0.001
Adjusted R^2	0.026	0.001
Residual Std. Error (df = 281555)	37.501	0.036
F Statistic (df = 7; 281555)	1,072.930***	25.486***
Note:	*p<0.1; **p<0.05; ***p<0.01	

Part B



The plot above illustrates the most significant relationship of part B. This is between time of year denoted by month and depature delay.

The regression results of Part B strived to determine the relationship of time of day, week and year between both delayed and cancelled flights. With each month relative to January as it is represented in the constant. Both month and hour are statistically significant in determining delayed depatures.The day of the week was not statistically significant in determining delayed depatures.In regards to months and cancellations all but 4 months were statistically significant.These months include March, August, October, and November. Also time of day and the day of the week were both statistically significant.In terms of magniture and statistical signifcance no month illustrated such a strong influence on depature delay as did july. On average holding all else constant if you were to fly in july you can expect departure delays to increase by 11.557 minutes. (reference to table 2)

Part C

```r
c1.reg<-lm(dep_delay~dest,a)
c2.reg<-lm(canceled~dest,a)
```

In Part C I attempted to find relationships between various destinations with both depature delays and cancelled flights. The number of destinations is very large so I decided to dicuss the destinations with statistically signifcant relationships with departure delays and cancellations (I have omitted this table from the pdf because it would have have taken up several pages in itself). I have instead depicted the models I

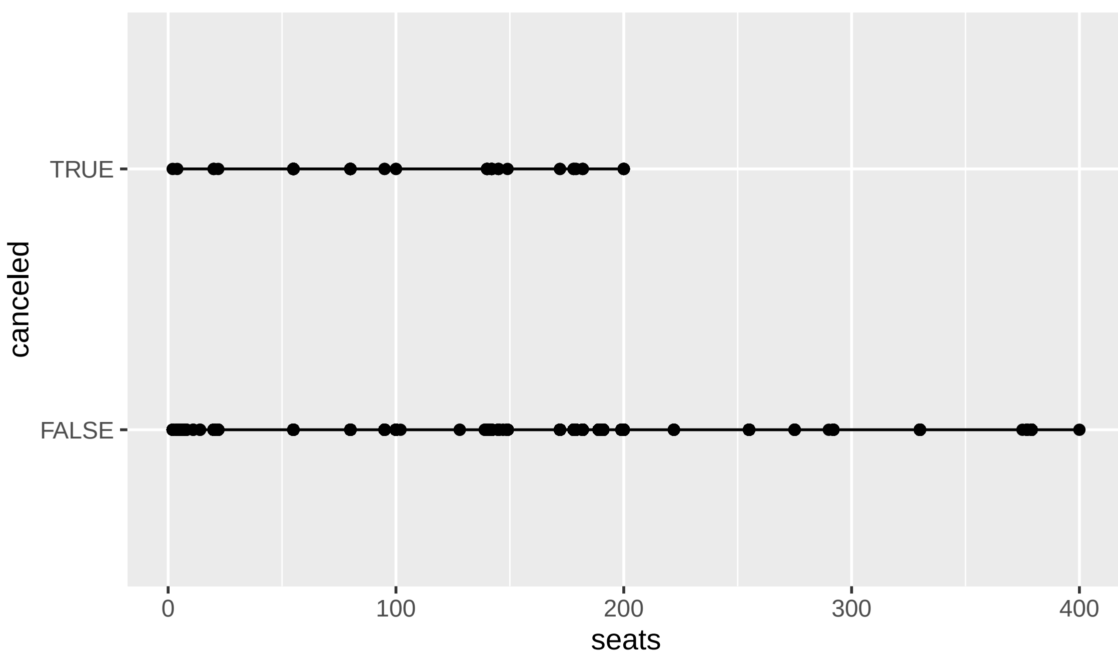
Table 2:

	Dependent variable:	
	dep_delay	canceled
	(1)	(2)
month2	0.765**	0.001**
	(0.344)	(0.0003)
month3	2.920***	0.0005
	(0.330)	(0.0003)
month4	3.772***	0.001***
	(0.331)	(0.0003)
month5	2.868***	0.001**
	(0.329)	(0.0003)
month6	10.732***	0.002***
	(0.332)	(0.0003)
month7	11.557***	0.003***
	(0.329)	(0.0003)
month8	2.514***	0.0001
	(0.328)	(0.0003)
month9	−3.028***	0.001***
	(0.333)	(0.0003)
month10	−3.625***	−0.0002
	(0.328)	(0.0003)
month11	−4.532***	0.0002
	(0.333)	(0.0003)
month12	6.311***	0.001**
	(0.333)	(0.0003)
day	−0.003	−0.00001*
	(0.008)	(0.00001)
hour	2.141***	0.0001***
	(0.014)	(0.00001)
Constant	−18.071***	−0.001*
	(0.321)	(0.0003)
Observations	328,521	328,521
R^2	0.084	0.001
Adjusted R^2	0.084	0.001
Residual Std. Error (df = 328507)	38.489	0.037
F Statistic (df = 13; 328507)	2,310.948***	18.741***
Note:	*p<0.1; **p<0.05; ***p<0.01	

created in order to determine the relationships.The most statistically significant relationships between destination and depature delays came from these destinations: BHM,CAE,DSM,OKC,RIC,TUL,and TYS. These were all statistically significant at a level greater than 99%. The most statistically significant relationships between destination and cancelled flights came from the following destinations: ALB, BDL, BGR, BHM, BNA, BWI, CAE, CHO, CHS, CLE, CMH, CVG, DAY, DCA, DSM, DTW, GRR,GSO, GSP, IAD, IND, JAC, LGA, MCI, MEM, MHT, MSP, OKC, ORD, ORF, PHL, PIT, PVD, RDU,RIC, SAV, SDF, SRQ, TUL and TYS. These destinations were all statistically significant at a level greater than 99%. In other words there were quite a bit more destinations that had a statistically signifcant relationship with cancelled flights.

(table ommitted)

Part D



The plot above illustrates the most significant relationship of part D. This is between the number of seats on the plane and cancelled flights.

Finally for Part D the regression model was designed to determine the effect of various plane characteristics on both depature delays and cancelled flights. Both the number of seats and engines were statistically significant in determining cancelled flights whereas only the number of seats were statistically significant in determining depature delays.

Overall the magnitude of the effects of the various plane characteristics are incredibly small and I believe are negligible in relative cancellations or depature delays. Also Engine type was not statistically signifcant in determining either depature delays or cancelled flights.

(reference to table 3)

Table 3:

	Dependent variable:	
	dep_delay	canceled
	(1)	(2)
engines	−0.370	0.016***
	(1.428)	(0.004)
seats	−0.031***	−0.0001***
	(0.001)	(0.00000)
engineReciprocating	−8.278	0.013
	(5.930)	(0.018)
engineTurbo-fan	−0.468	−0.0002
	(6.022)	(0.019)
engineTurbo-jet	−2.024	−0.008
	(6.028)	(0.019)
engineTurbo-prop	−3.746	−0.015
	(8.480)	(0.026)
engineTurbo-shaft	−4.063	−0.004
	(6.193)	(0.019)
Constant	18.863***	0.005
	(6.019)	(0.019)
Observations	279,971	284,170
R^2	0.003	0.008
Adjusted R^2	0.003	0.008
Residual Std. Error	40.510 (df = 279963)	0.125 (df = 284162)
F Statistic	135.281*** (df = 7; 279963)	344.382*** (df = 7; 284162)
Note:		*p<0.1; **p<0.05; ***p<0.01