

STAT495 HW#3

YOUR NAME HERE

September 24, 2015

1: airline delays

Modify the sample code to access the airline delays data to answer the following question: what airports and airlines in the United States ran into trouble on January 26, 2015? How do those delays compare to the previous week (January 19, 2015)?

Your report should be no more than 2 pages of text with up to 2 additional tables and 3 figures. Please pay particular attention to formatting of figures and tables.

```
library(mosaic)
library(RMySQL)
```

```
## Loading required package: DBI
```

```
db <- src_mysql(host = "mysql-research.amherst.edu",
               user = "math230", password = "math230pass",
               dbname = "math230")
```

```
airports <- tbl(db, "airports")
planes <- tbl(db, "planes")
carriers <- tbl(db, "carriers")
flights <- tbl(db, "flights")
```

```
aprilfools <-
  flights %>%
  filter(month==1, day==1, year==2015)
aprilfools <- data.frame(aprilfools)
aprilfools <- aprilfools %>%
  mutate(realdelay = ifelse(is.na(arr_delay), 240, arr_delay))
favstats(~ arr_delay, data=aprilfools)
```

```
##   min   Q1 median  Q3   max     mean    sd    n missing
##  -54  -12     -3  12 1201  5.352496 34.77 13464     486
```

```
sort(mean(cancelled ~ origin, data=aprilfools), decreasing=TRUE) %>% head(10)
```

```
##      ABI      HOB      ROW      SJT      MAF      SPS      TXK
## 1.0000000 1.0000000 1.0000000 1.0000000 0.7500000 0.7500000 0.6666667
##      LAW      APN      CMX
## 0.6000000 0.5000000 0.5000000
```

```
save(aprilfools, file="APRILFOOLS.Rda") # load this in a different Rmd file to run your analyses
```

2: snowstorm insurance

A company buys a policy to insure its revenue in the event of major snowstorms that shut down business. The policy pays nothing for the first such snowstorm of the year and 10,000 for each one thereafter, until the end of the year. The number of major snowstorms per year that shut down business is assumed to have a Poisson distribution with mean 1.5. Find the analytic solution as well as an empirical estimate (be sure to use at least 10,000 simulations and provide a 99% confidence interval for your estimate).

3: ISL exercise 3.15 (page 126)

for part d) please be sure to comment on the results (in a single table).

```
Boston <- MASS::Boston
names(Boston)
```

```
## [1] "crim"    "zn"      "indus"   "chas"    "nox"     "rm"      "age"
## [8] "dis"     "rad"     "tax"     "ptratio" "black"   "lstat"   "medv"
```

Please pay particular attention to formatting. See the following code chunks for examples (and see <https://cran.r-project.org/web/packages/xtable/vignettes/xtableGallery.pdf>).

```
x <- rnorm(100)
y <- 0.5 + 3*x + rnorm(100)
mod <- lm(y ~ x)
require(xtable)
```

```
## Loading required package: xtable
```

```
displaymod <- xtable(mod)
digits(displaymod) <- 3
```

```
displaymod
```

% latex table generated in R 3.2.1 by xtable 1.7-4 package % Fri Sep 25 09:15:37 2015

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.428	0.099	4.337	0.000
x	3.016	0.101	29.968	0.000

```
options(xtable.comment = FALSE)
table <- cbind(result1=c("Yes", "No"), result2=c("Maybe", "Yes"))
xtable(table)
```

	result1	result2
1	Yes	Maybe
2	No	Yes