# Untitled

*Honey Berk*

*March 16, 2017*

```
library(data.table)
library(dplyr)
```

```
## ------------------------------------------------------------------------
```

```
## data.table + dplyr code now lives in dtplyr.
## Please library(dtplyr)!
```

```
## ------------------------------------------------------------------------
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:data.table':
##
##     between, first, last
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(ggplot2)


# Load data

post.attitude <- fread(input = "./REVISED_FINAL_POST_ATTITUDE_rename.csv",
                       header = TRUE, stringsAsFactors = TRUE)
Pre_post <- fread(input = "./Pre_post.csv",
                  header = TRUE, stringsAsFactors = TRUE)

### linear Regression on A0.3, A0.4, A0.57,A0.58,Experience
#### not enough variaion in 0.5, 0.60

# #check levels
# sapply(post.attitude %>%
#          select(A0.3,A0.4,A0.56,A0.57,A0.58,
#                 A0.67, A0.68, A0.69, A0.70, A0.71, A0.72,
#                 Experience),function(x)levels(x))

# merge pre_post and post.attitude
pre_post.attitude <- merge(Pre_post,post.attitude,by='BPL.BLD.ID', all.x = TRUE)

# select E group
pre_post.attitude_E <- pre_post.attitude %>%
  filter(LogCheck == "Y")
```

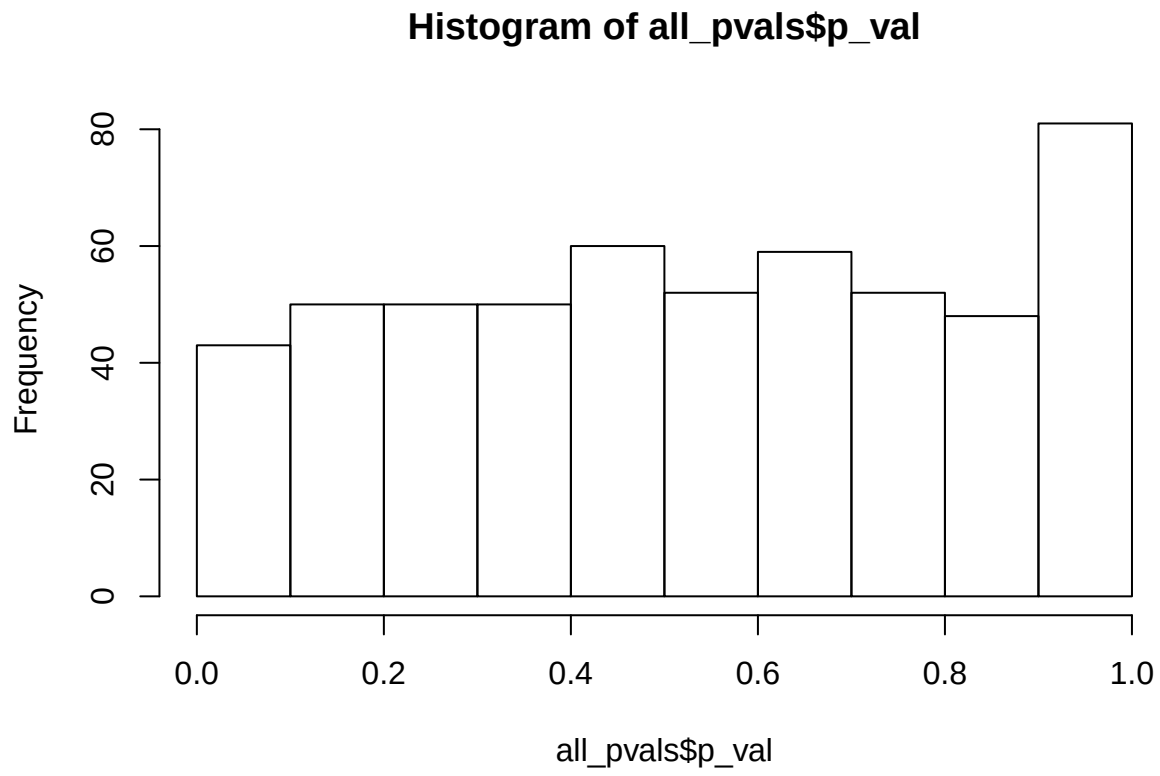
```r
# ONE - NO TECH, NO AGE OR EXPERIENCE
all_pvals <- data.frame(varname=as.character(), col=as.character(), p_val=as.numeric())

for (col in names(pre_post.attitude_E)[grep("X", names(pre_post.attitude_E))]) {
  # cat(col, ":\n")
  smy <- summary(lm(paste(col, "~ A0.3 + A0.4 + A0.56 + A0.57 + A0.58 + A0.59 + Age + Experience + A0.6
                    data = pre_post.attitude_E))

  all_pvals <- rbind(all_pvals,
                     data.frame(varname = names(smy$coefficients[,4]), col = col, p_val = smy$coeffic

}
```

```r
hist(all_pvals$p_val)
```




```r
all_pvals %>% filter(p_val < .05)
```

```
##              varname  col       p_val
## 1           A0.57Yes X1.4 0.017476011
## 2           A0.57Yes X2.1 0.030605236
## 3        (Intercept) X3.1 0.041709267
## 4           Age44-54 X3.1 0.004290497
## 5             Age55+ X3.1 0.011591706
## 6   Experience15-20Y X3.1 0.006519213
## 7     Experience20+Y X3.1 0.002421040
## 8    Experience5-10Y X3.1 0.003974522
```

```
## 9           A0.56Yes X3.3 0.035350768
## 10          A0.57Yes X3.3 0.001050115
## 11          A0.59Yes X3.3 0.027619137
## 12          A0.56Yes X3.6 0.040277584
## 13          A0.56Yes X4.5 0.029526885
## 14          A0.58Yes X4.5 0.048757364
## 15          Age44-54 X4.5 0.009357408
## 16            Age55+ X4.5 0.035390824
## 17  Experience15-20Y X4.5 0.008122051
## 18    Experience20+Y X4.5 0.018161658
## 19          A0.69Yes X4.8 0.023024443
## 20          A0.57Yes X6.2 0.038793549
## 21   Experience5-10Y X6.3 0.039922465
## 22 A0.3Very involved X7.1 0.035656837
## 23          A0.56Yes X7.1 0.049678794
## 24            Age55+ X7.1 0.013745740
## 25  Experience15-20Y X7.1 0.029065456
## 26    Experience20+Y X7.1 0.013113864
```