Enron Data for EDA

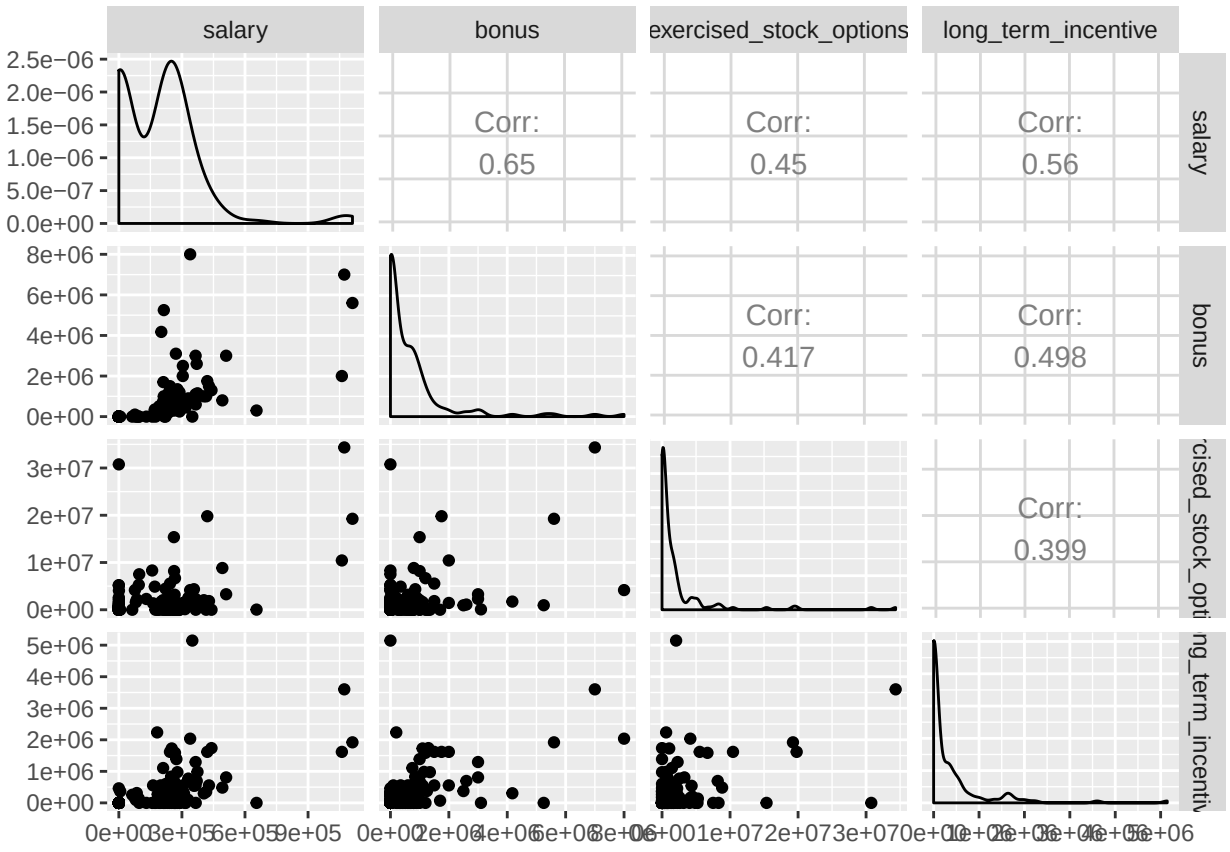
R. Lorenzo

February 13, 2017

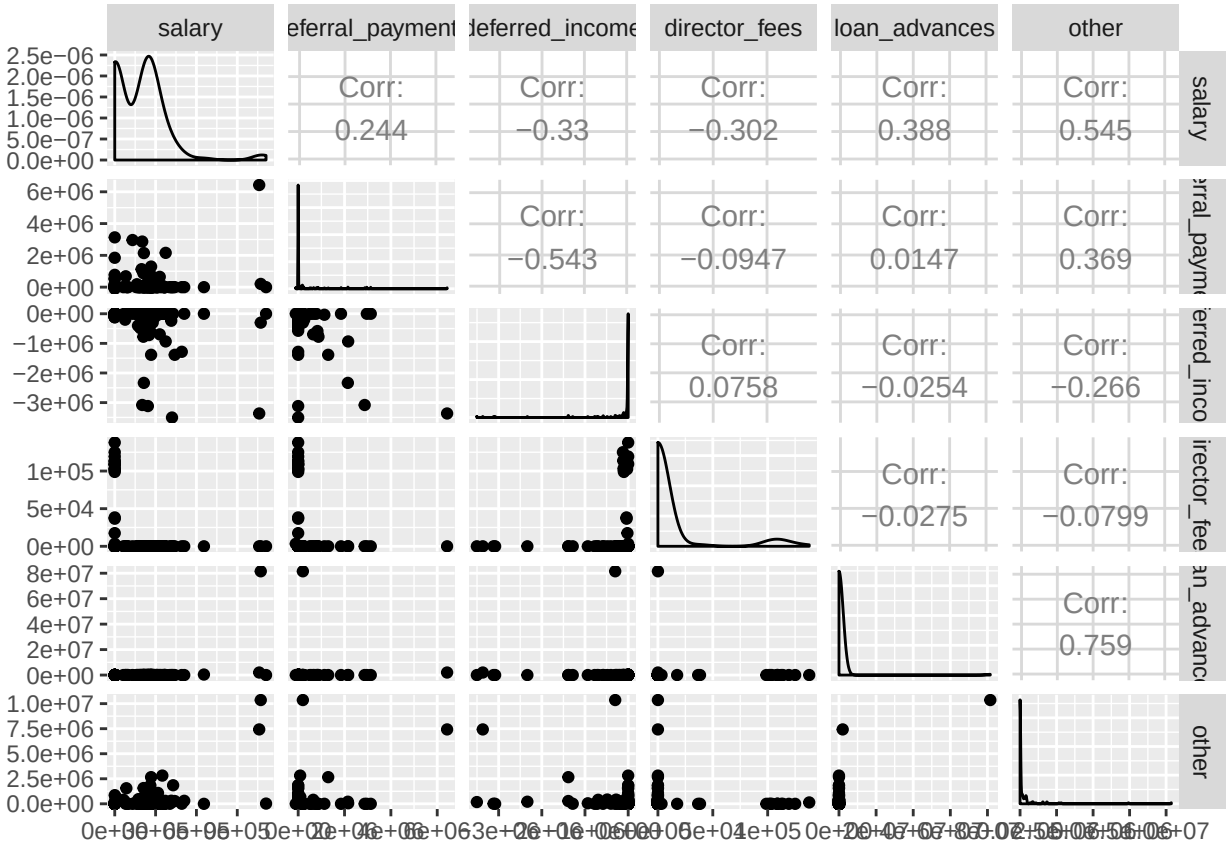
```
library(ggplot2)
library(tidyr)
library(alr3)
library(gridExtra)
library(GGally)
library(memisc)
library(reshape2)
library(readr)
library(memisc)
library(dtplyr)
setwd("C:/Users/rl1891/version-control/MachineLearning/ud120-projects/final_project")
```

```
library(readr)
enron_for_eda <- read_csv("C:/Users/rl1891/version-control/MachineLearning/ud120-projects/final_project/
    col_types = cols(poi = col_factor(levels = c("False",
        "True"))))
View(enron_for_eda)
```

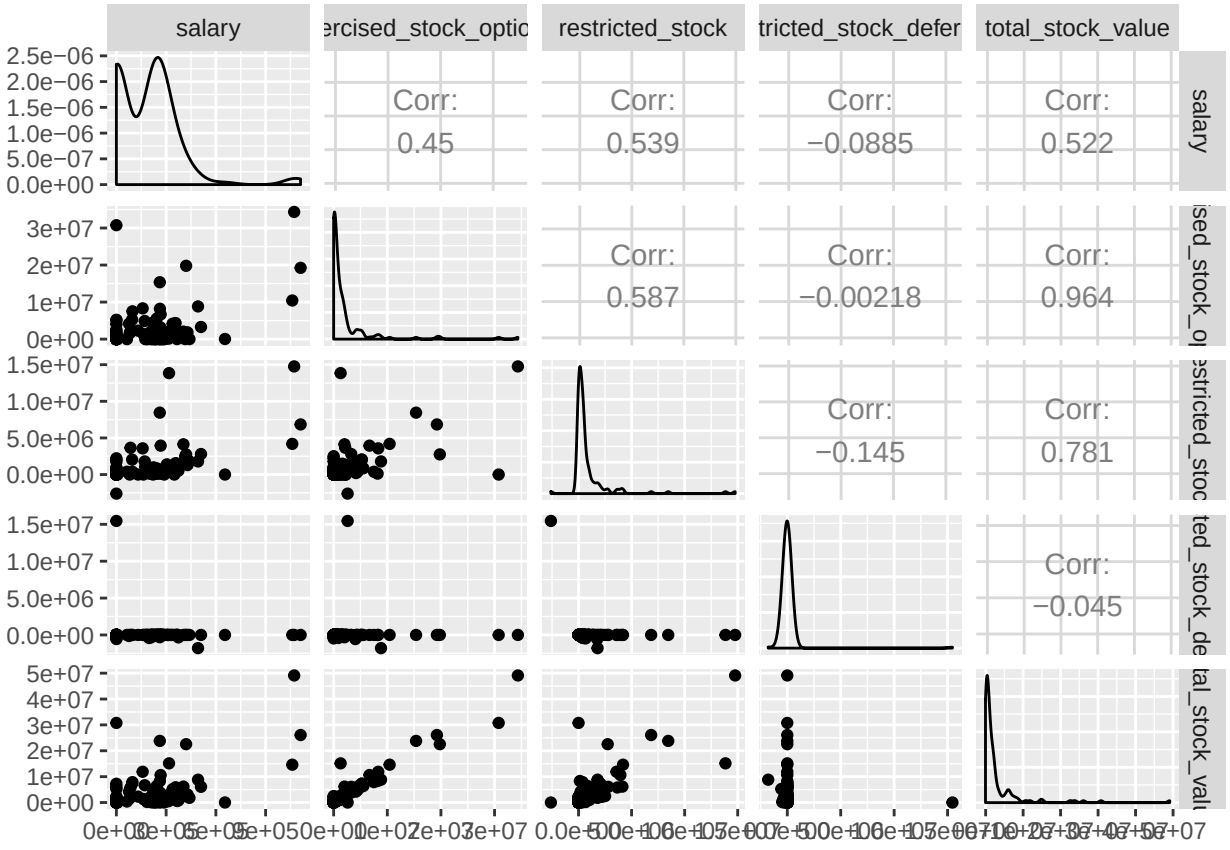
```
ggpairs(enron_for_eda, columns = c("salary", "bonus", "exercised_stock_options",
                                   "long_term_incentive"))
```



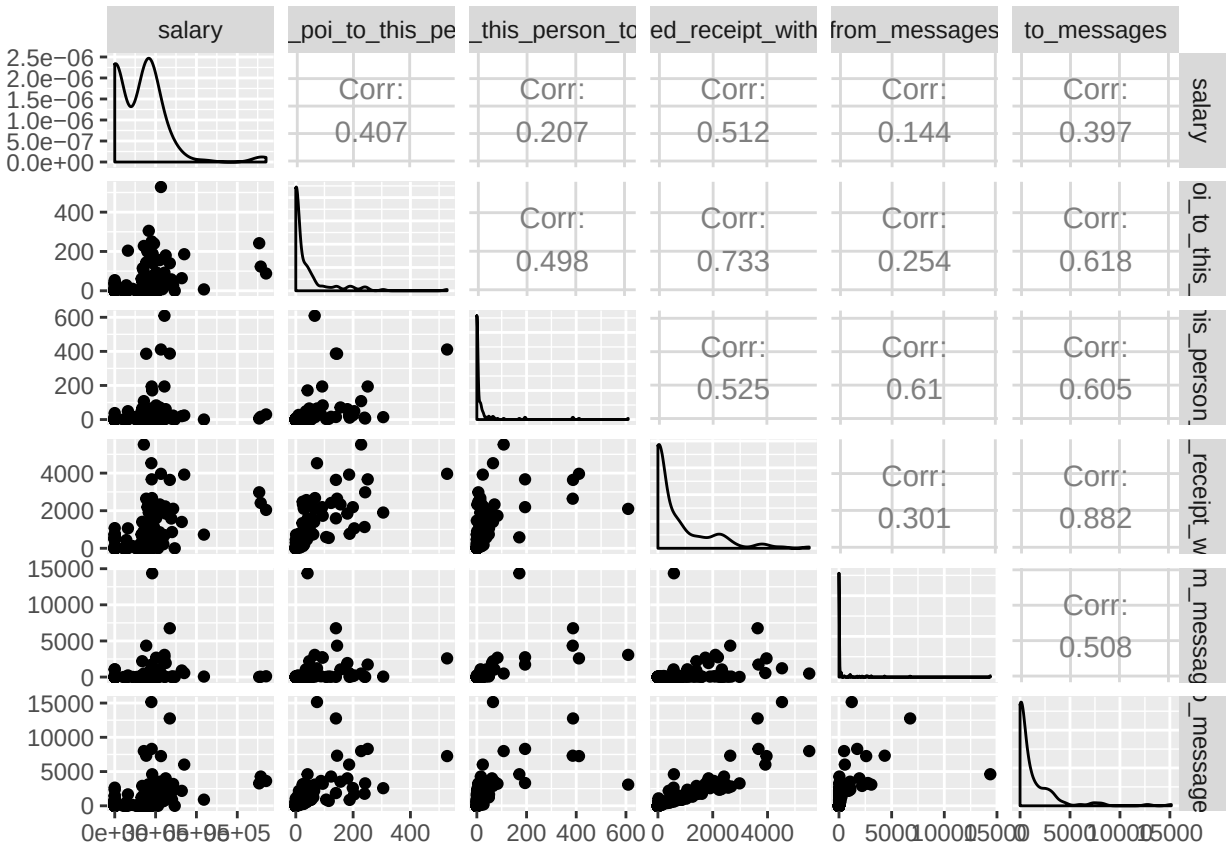

```
ggpairs(enron_for_eda, columns = c("salary", "deferral_payments",
                                   "deferred_income", "director_fees",
                                   "loan_advances", "other"))
```



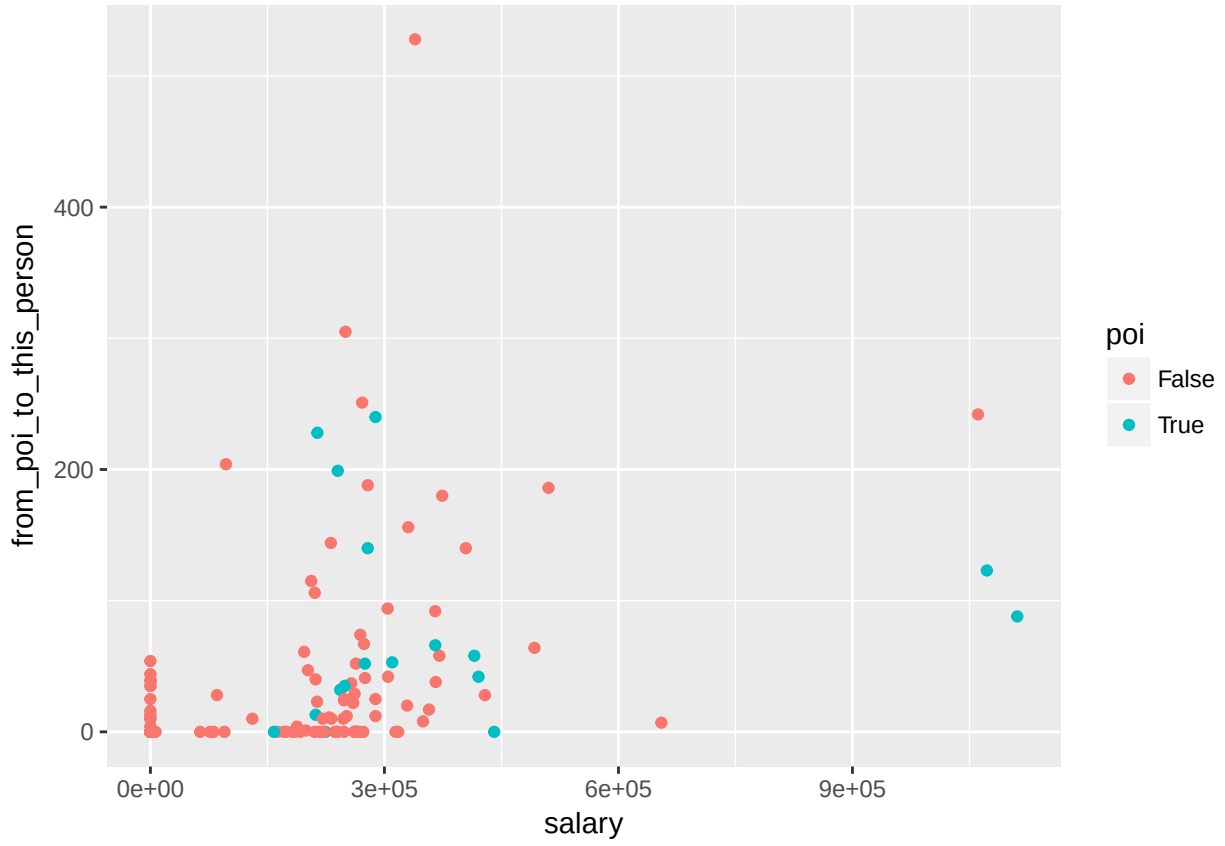

```
ggpairs(enron_for_eda, columns = c("salary", "exercised_stock_options",
                                   "restricted_stock", "restricted_stock_deferred",
                                   "total_stock_value"))
```



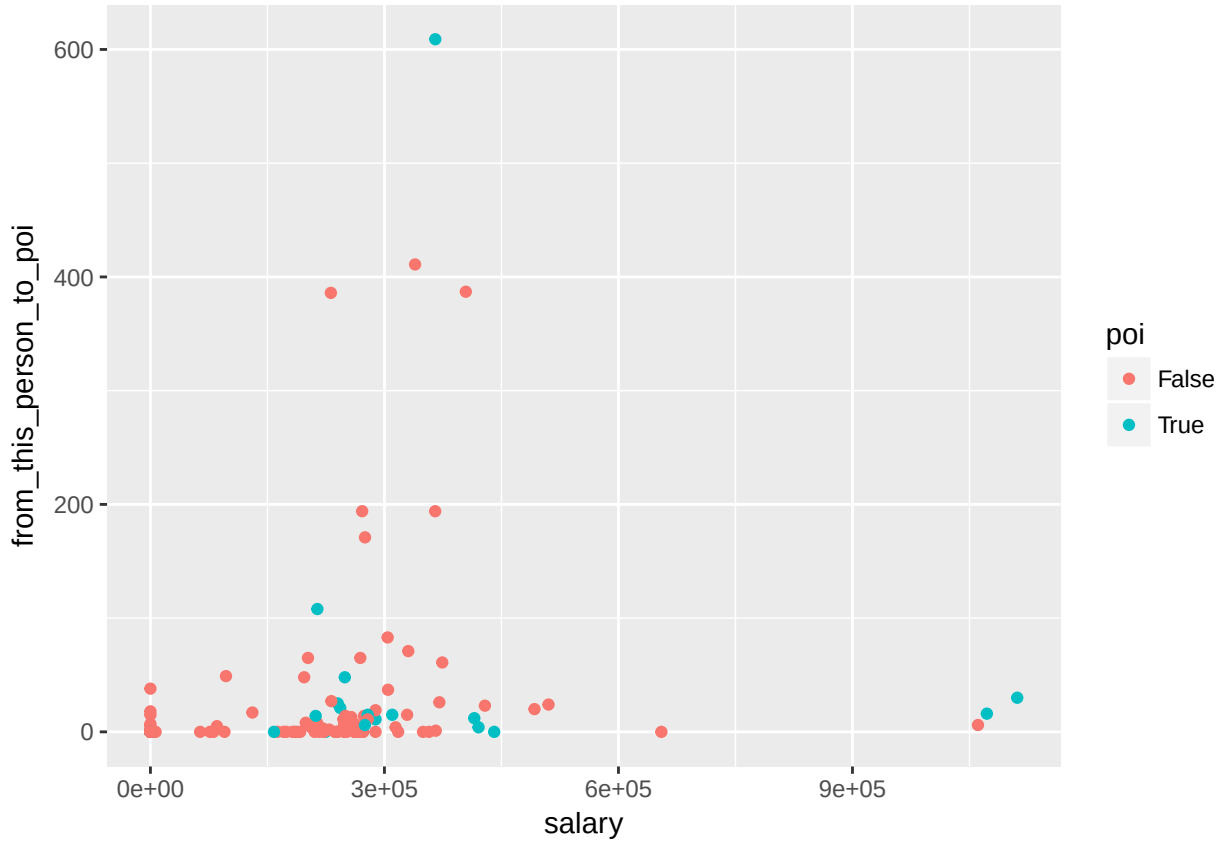

```
ggpairs(enron_for_eda, columns = c("salary", "from_poi_to_this_person",
                                   "from_this_person_to_poi", "shared_receipt_with_poi",
                                   "from_messages", "to_messages"))
```



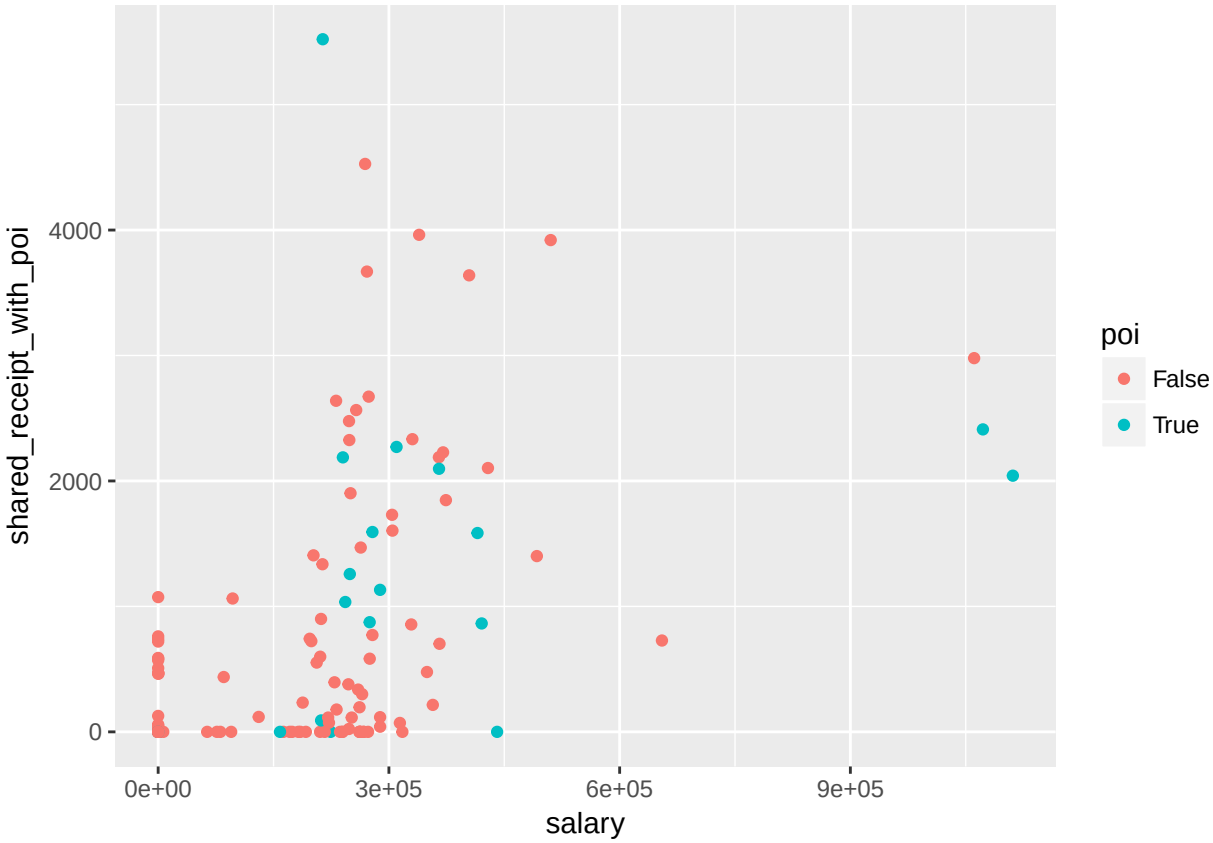

```
ggplot(aes(x=salary, y=from_poi_to_this_person, color = poi), data = enron_for_eda) +
  geom_point()
```




```
ggplot(aes(x=salary, y=from_this_person_to_poi, color = poi), data = enron_for_eda) +
  geom_point()
```




```
ggplot(aes(x=salary, y=shared_receipt_with_poi, color = poi), data = enron_for_eda) +
  geom_point()
```


shared_receipt_with_poi
4000
2000
0
0e+00
3e+05
6e+05
9e+05
salary
poi
False
True