Final Project STA130

Ralisha Woodhouse

2022-11-14

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.4.0      v purrr   0.3.5
## v tibble  3.1.8      v dplyr   1.0.10
## v tidyr   1.2.1      v stringr 1.4.1
## v readr   2.1.3      v forcats 0.5.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
df_tidy<-read_csv("Tidy_DF.csv")
```

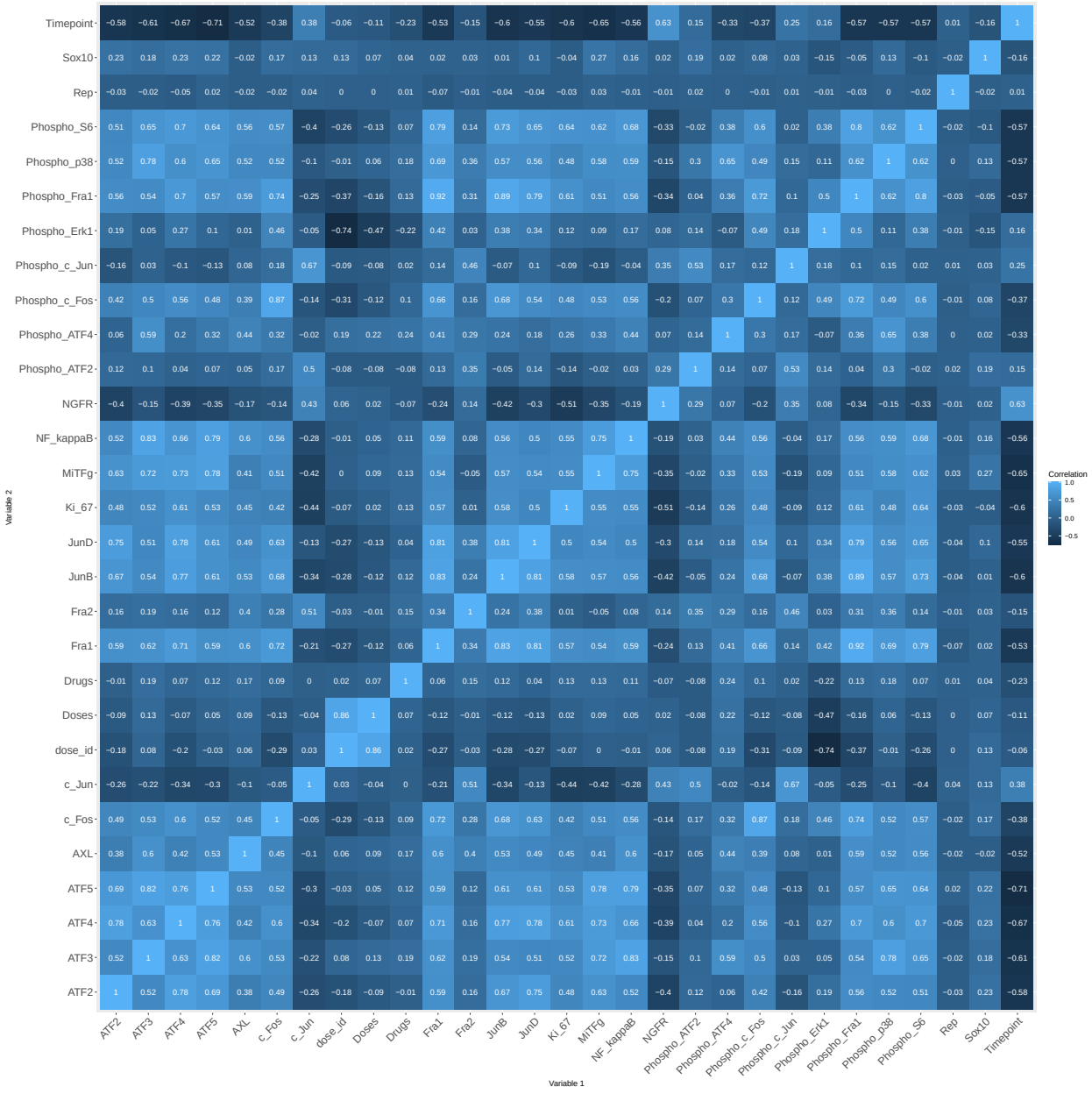
```
## Rows: 540792 Columns: 30
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## dbl (30): ...1, Phospho_c_Fos, Phospho_c_Jun, Phospho_ATF2, Phospho_Fra1, c_...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
glimpse(df_tidy)
```

```
## Rows: 540,792
## Columns: 30
## $ ...1          <dbl> 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16~
## $ Phospho_c_Fos <dbl> 2.547698, 2.657264, 2.566059, 2.351795, 2.675345, 2.4033~
## $ Phospho_c_Jun <dbl> 2.426777, 2.621776, 2.602058, 2.456012, 2.614406, 2.4563~
## $ Phospho_ATF2  <dbl> 2.354171, 2.381176, 2.384345, 2.092770, 2.658403, 2.3971~
## $ Phospho_Fra1  <dbl> 3.648492, 4.072734, 3.844549, 3.422671, 3.732772, 3.4572~
## $ c_Fos         <dbl> 2.851415, 2.983508, 2.913421, 2.820291, 2.961870, 2.5836~
## $ c_Jun         <dbl> 2.366601, 2.737159, 2.756987, 2.284825, 2.604511, 2.6345~
## $ Fra1          <dbl> 3.796645, 3.995531, 3.960424, 3.648597, 3.884726, 3.6074~
## $ JunD          <dbl> 3.300881, 3.336940, 3.394186, 3.246594, 3.426226, 3.1894~
## $ ATF2          <dbl> 3.076601, 2.968231, 3.113777, 2.927099, 3.097459, 2.9632~
## $ JunB          <dbl> 3.524831, 3.728830, 3.496554, 3.364566, 3.723072, 3.4038~
## $ Fra2          <dbl> 2.425722, 2.953737, 2.688302, 2.518786, 3.127684, 2.6986~
## $ ATF4          <dbl> 2.835745, 2.973435, 2.827416, 2.778560, 2.866927, 2.6434~
## $ Phospho_ATF4  <dbl> 3.574612, 3.506133, 3.797322, 3.431327, 3.617248, 3.4802~
## $ Phospho_Erk1  <dbl> 3.874797, 3.915648, 3.911892, 3.562602, 3.910799, 3.7781~
```

```
## $ NGFR          <dbl> 3.128172, 3.105526, 3.146136, 2.968830, 3.136196, 3.1266~
## $ Phospho_S6    <dbl> 4.468469, 4.227182, 4.367623, 4.048401, 4.021664, 4.1744~
## $ ATF3          <dbl> 3.683911, 3.697249, 3.795028, 3.567414, 3.661982, 3.5805~
## $ ATF5          <dbl> 3.083728, 3.211368, 3.324015, 3.105321, 3.102642, 3.0766~
## $ Phospho_p38   <dbl> 2.909924, 3.032803, 3.015247, 2.849349, 3.170207, 2.9306~
## $ Ki_67         <dbl> 4.053889, 4.121248, 4.086724, 3.705978, 3.679998, 3.5542~
## $ NF_kappaB     <dbl> 3.770033, 3.641639, 3.937815, 3.548643, 3.687348, 3.6468~
## $ MiTFg         <dbl> 3.974078, 3.595271, 4.097996, 3.725713, 3.805543, 3.5868~
## $ AXL           <dbl> 3.536432, 3.732794, 3.609001, 3.223876, 3.600571, 3.5281~
## $ Sox10         <dbl> 3.686878, 3.668114, 3.781692, 3.700308, 3.755307, 3.5294~
## $ Timepoint     <dbl> 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0~
## $ Drugs         <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,~
## $ dose_id       <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,~
## $ Doses         <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,~
## $ Rep           <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,~
```

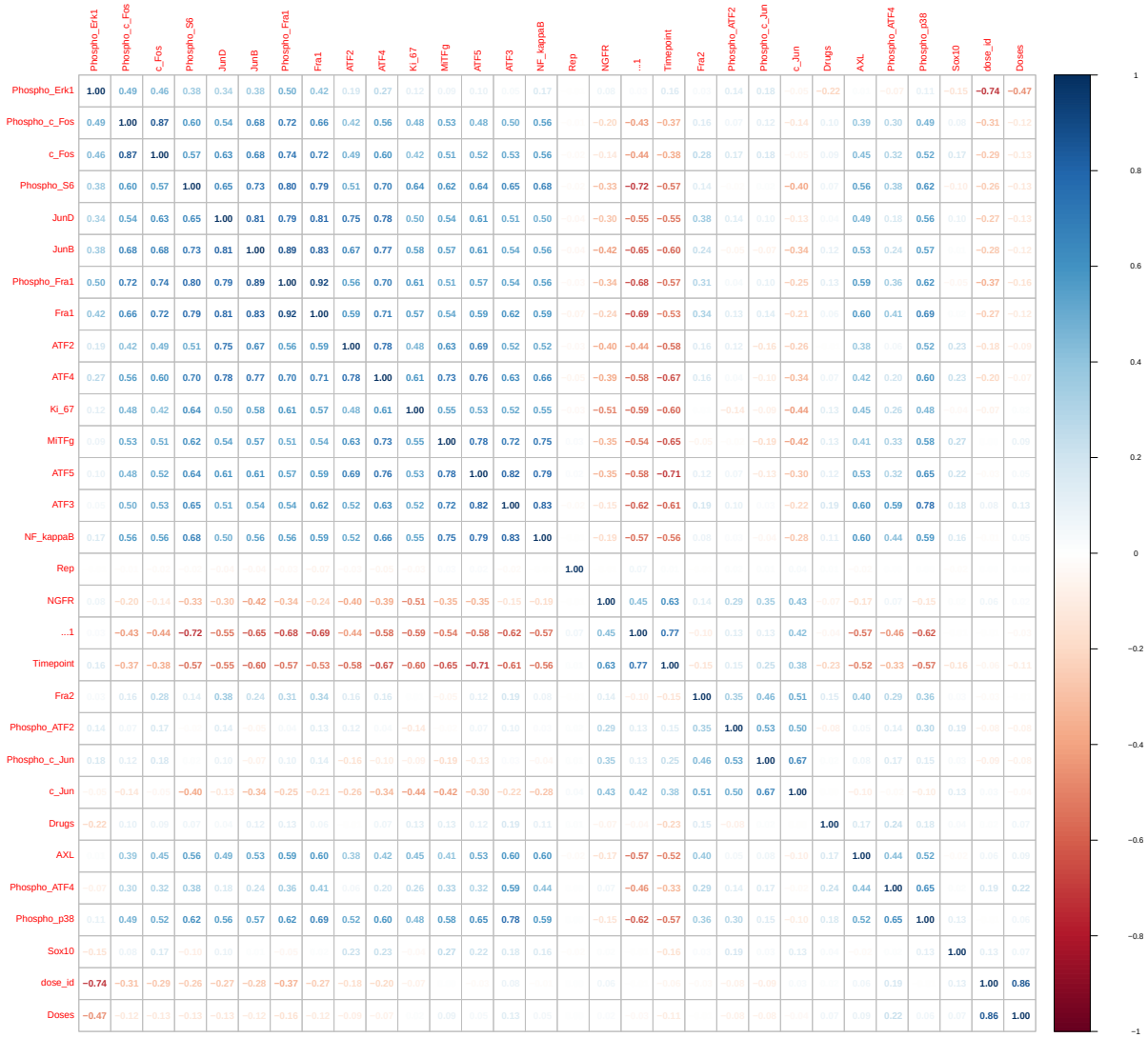
```
df_tidy %>% select(-...1) %>%
cor() %>% as_tibble(rownames="rowname") %>%
pivot_longer(cols=!rowname,
names_to="Variable 1",
values_to="Correlation") %>%
rename("Variable 2"=rowname) %>%
ggplot(aes(x=`Variable 1`, y=`Variable 2`,
fill=Correlation,
label=round(Correlation,2))) +
geom_tile() + geom_text(color="white") +
theme(axis.text.x = element_text(angle = 45, hjust=1), axis.text = element_text(size = 15))
```




```
#This is just another method for doing the correlation matrix first --> install.packages("corrplot")
library(corrplot)
```

```
## corrplot 0.92 loaded
```

```
Corr<-cor(df_tidy)
corrplot(Corr,order="hclust", method="number")
```




```
#This one is nicer because it orders the correlations (the positive correlations (blue) are top left).
```