````
---
title: "R Notebook"
output: html_notebook
---

This is an [R Markdown](http://rmarkdown.rstudio.com) Notebook. When you execute code within the notebook, the results appear beneath the code.

Try executing this chunk by clicking the *Run* button within the chunk or by placing your cursor inside it and pressing *Ctrl+Shift+Enter*.

```{r ANSWER TO Q1}
#commented as it asks for restarting R
#install.packages("remotes")
#remotes::install_github("Shreya-Vaish/BSE658") #saves them to temp, how can i check?
library(tidyverse)
library(dplyr)
library(tibble)
```

Add a new chunk by clicking the *Insert Chunk* button on the toolbar or by pressing *Ctrl+Alt+I*.

When you save the notebook, an HTML file containing the code and output will be saved alongside it (click the *Preview* button or press *Ctrl+Shift+K* to preview the HTML file).

The preview shows you a rendered HTML copy of the contents of the editor. Consequently, unlike *Knit*, *Preview* does not run any R code chunks. Instead, the output of the chunk when it was last run in the editor is displayed.



```{r}
getwd()
smokeban<-read_csv("C:\\Users\\ushad\\Downloads\\SmokeBan.csv") #use double backslash or forward slashes
```
```{r}
spec(smokeban)
```
```{r}
#?cols_condense
#?spec
smokeban #im ok with the coltype being displayed so I am not going to make its display false or condense it
```
```{r}
sum(is.na(smokeban$hispanic))
filter(smokeban,hispanic=="yes" & smoker=="yes")
youngsmokers<-nrow(filter(smokeban, age<25))
newnum<-youngsmokers/(nrow(smokeban))
newnum*100 #how to write this into a string?
```
```{r}
#select(smokeban, c(afam,hispanic))
select(smokeban, -c(afam,hispanic))
```
```{r}
newsel<-select(smokeban,smoker, age:afam) #afam means african american
newsel
```

#Q2 answer was given already in q file
````

```{r ANSWER TO Q3}
rename(smokeban, idk=afam)
smokeban
#rename only shows a temporary change, in order to save it, you need to assign it the the file like smokeban<-rename(smokeban, idk=afam or assign to a new variable)
```
```{r}
getwd()
phddata<-read_csv("C:\\Users\\ushad\\Downloads\\PhDPublications.csv")
phddata
newphddata<-mutate(phddata,success_dependency = prestige/mentor) #mutate does not work unless you assign a vairable to the changed dataframe
arrange(newphddata,success_dependency)
#prestige is for the grduate program and mentor means number of papers a mentor has published
```
```{r}
arrange(newphddata,desc(success_dependency)) #there are infinite values because some mentors have 0 published papers!
```
```{r}
write.csv(newphddata, file="newphddata.csv")
list.files()
getwd()
read.csv(file="C:/Users/ushad/Documents/GitHub/BSE658/Module 2/newphddata.csv")
```


```{r ANSWER TO Q6}
shipdata<-read_csv(file="C:/Users/ushad/Downloads/ShipAccidents.csv")
shipdata
shipdata %>% select(-incidents) %>% arrange(desc(type)) %>% arrange(construction) #it first arranged wrt to construction, then type
modifiedshipdata <- shipdata %>% select(-incidents) %>% arrange(construction)  %>% arrange(desc(type))
modifiedshipdata
```
```{r ANSWER TO Q4}
#another way of adding columns
shipdata %>% add_column(newcol = shipdata$service - shipdata$incidents, newcol2='0')
#i think mutate has some functions like cummean() associated with itself which add_column does not, which makes mutate more useful
```