title: “session 2” author: “venesha” date: “1/13/2020” output: html\_document editor\_options: chunk output:console

```r  
library(tidyverse)

## -- Attaching packages -------------------------------------------- tidyverse 1.3.0 --

## v ggplot2 3.2.1 v purrr 0.3.3  
## v tibble 2.1.3 v dplyr 0.8.3  
## v tidyr 1.0.0 v stringr 1.4.0  
## v readr 1.3.1 v forcats 0.4.0

## -- Conflicts ----------------------------------------------- tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

mobile<-read.csv("https://raw.githubusercontent.com/vmandela99/Class-assignment-12th-july-2019/master/mobilemoney\_data.csv")

## report on summary

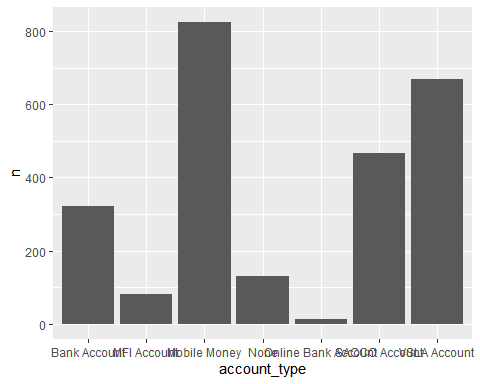
skimr::skim(mobile)

The data has 2509 number of rows and 28 number of columns.

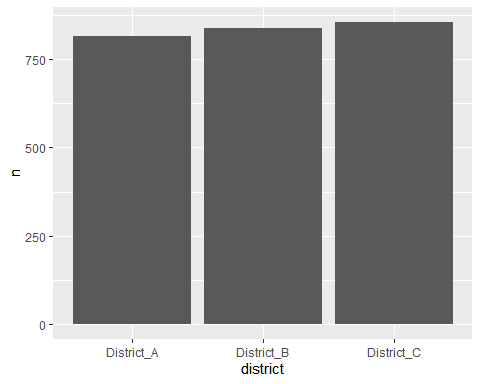
# how many account types are there  
mobile %>% count(account\_type,sort=TRUE)

## # A tibble: 7 x 2  
## account\_type n  
## <fct> <int>  
## 1 Mobile Money 825  
## 2 VSLA Account 669  
## 3 SACCO Account 467  
## 4 Bank Account 323  
## 5 None 131  
## 6 MFI Account 82  
## 7 Online Bank Account 12

mobile %>% count(account\_type) %>% ggplot(aes(x=account\_type,y=n)) + geom\_col()

 There are seven account types in the mobile data

# which districts has many people  
mobile %>% count(district) %>% ggplot(aes(x=district,y=n)) + geom\_col()

 District\_c has the highest number of people which is 856

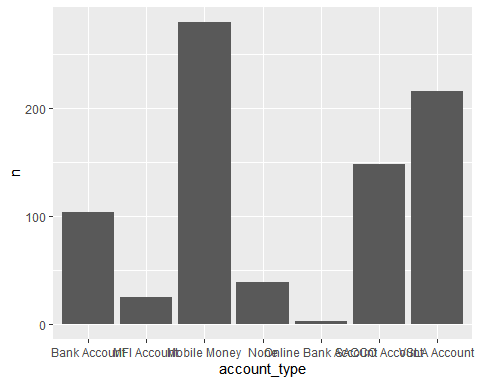
# what percentage of people from district A has sacco accounts  
mobile %>% count(district, account\_type) %>% mutate(percent=n/sum(n)\*100)

## # A tibble: 21 x 4  
## district account\_type n percent  
## <fct> <fct> <int> <dbl>  
## 1 District\_A Bank Account 104 4.15   
## 2 District\_A MFI Account 25 0.996  
## 3 District\_A Mobile Money 280 11.2   
## 4 District\_A None 39 1.55   
## 5 District\_A Online Bank Account 3 0.120  
## 6 District\_A SACCO Account 148 5.90   
## 7 District\_A VSLA Account 216 8.61   
## 8 District\_B Bank Account 118 4.70   
## 9 District\_B MFI Account 27 1.08   
## 10 District\_B Mobile Money 268 10.7   
## # ... with 11 more rows

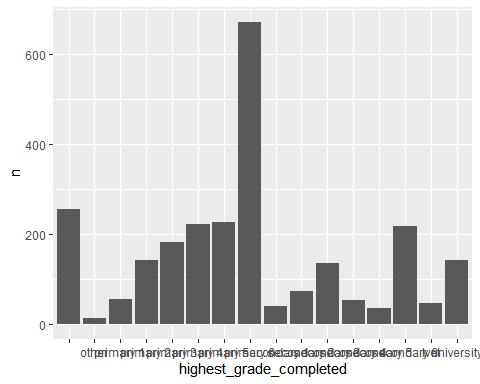
mobile %>% filter(district == "District\_A") %>% count(account\_type) %>% mutate(percent=n/sum(n)\*100)

## # A tibble: 7 x 3  
## account\_type n percent  
## <fct> <int> <dbl>  
## 1 Bank Account 104 12.8   
## 2 MFI Account 25 3.07   
## 3 Mobile Money 280 34.4   
## 4 None 39 4.79   
## 5 Online Bank Account 3 0.368  
## 6 SACCO Account 148 18.2   
## 7 VSLA Account 216 26.5

mobile %>% filter(district == "District\_A") %>% count(account\_type) %>% mutate(percent=n/sum(n)\*100) %>% ggplot(aes(x=account\_type,y=n)) + geom\_col()

 5.90 is the percentage of people from District A with sacco accounts

# how many grades completed are there?  
mobile %>% count(highest\_grade\_completed) %>% ggplot(aes(x=highest\_grade\_completed,y=n)) + geom\_col()

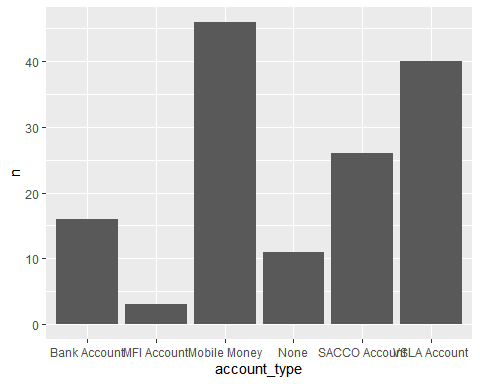
 There are 16 grades completed

# what is the largest weight in the survey  
max(mobile$weight)

## [1] 4812.165

The largest weight in the survey is 4812’165

# what percentage of people are in university and have mobile money account?  
mobile %>% filter(highest\_grade\_completed == "university") %>% count(account\_type) %>% mutate(percent=n/sum(n)\*100) %>% ggplot(aes(x=account\_type,y=n)) + geom\_col()

 There are 32.4 percent people in the university with mobile money