VisualizationExploration

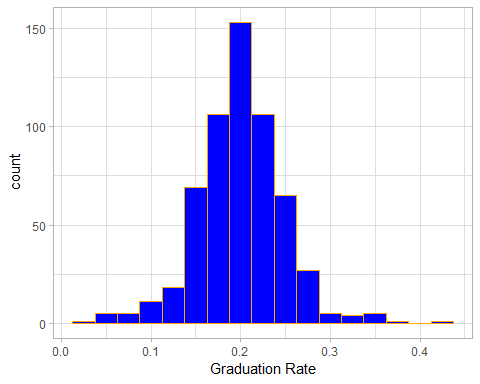
Kevin Letourneau

May 22, 2019

data <- read.csv(file="C:/Users/fuses/Downloads/institution\_data\_R.csv",header = TRUE)  
head(data)

## ï..institution\_name  
## 1 Alabama A&M University  
## 2 University of Alabama at Birmingham  
## 3 University of Alabama at Huntsville  
## 4 Alabama State University  
## 5 Auburn University at Montgomery  
## 6 Auburn University  
## basic grad\_rate  
## 1 Masters Colleges and Universities--larger programs 0.142  
## 2 Research Universities--very high research activity 0.209  
## 3 Research Universities--very high research activity 0.209  
## 4 Masters Colleges and Universities--larger programs 0.116  
## 5 Masters Colleges and Universities--larger programs 0.154  
## 6 Research Universities--high research activity 0.215  
## student\_count spending\_per\_award full\_time\_pct full\_time\_count  
## 1 4051 105331 0.938 3906  
## 2 11502 136546 0.727 10032  
## 3 5696 64418 0.744 5000  
## 4 5356 132407 0.910 5035  
## 5 4322 58541 0.694 3571  
## 6 19799 71999 0.910 19635  
## med\_sat\_value aid\_value endow\_value grad\_on\_time\_pct pell\_value  
## 1 823 7142 3808 0.10 0.712  
## 2 1146 6088 24136 0.29 0.351  
## 3 1180 6647 11502 0.16 0.328  
## 4 830 7256 13202 0.08 0.827  
## 5 970 4327 10736 0.09 0.401  
## 6 1215 8875 22092 0.37 0.169  
## fresh\_retain\_value full\_time\_fac\_pct EndowXSpend PellXSat RetainXSat  
## 1 0.631 0.828 401100448 585.976 519.313  
## 2 0.802 0.924 3295674256 402.246 919.092  
## 3 0.810 0.655 740935836 387.040 955.800  
## 4 0.622 0.670 1748037214 686.410 516.260  
## 5 0.632 0.601 628496176 388.970 613.040  
## 6 0.895 0.887 1590601908 205.335 1087.425  
## AidXSat AidXEndow id  
## 1 5877866 27196736 1  
## 2 6976848 146939968 2  
## 3 7843460 76453794 3  
## 4 6022480 95793712 4  
## 5 4197190 46454672 5  
## 6 10783125 196066500 6

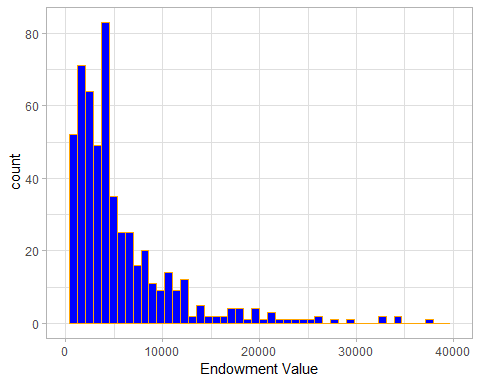
#install.packages("tidyverse")  
#install.packages("gcookbook")  
#install.packages("MASS")  
#install.packages("ggplot2")  
#install.packages("dplyr")  
  
  
  
grad <- ggplot(data = data, aes(x=grad\_rate)) +  
 geom\_histogram(binwidth = 0.025, bins = 20, fill = "blue", color = "orange")+  
 labs(x="Graduation Rate")+  
 theme\_light()  
  
grad



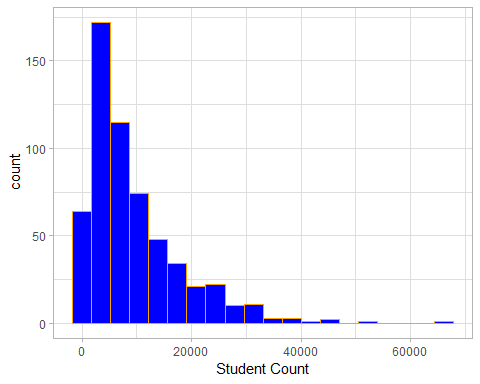
endow <- ggplot(data = data, aes(x=endow\_value)) +  
 geom\_histogram(bins = 50, fill = "blue", color = "orange")+  
 labs(x="Endowment Value")+  
 xlim(0,40000)+  
 theme\_light()  
  
endow

## Warning: Removed 11 rows containing non-finite values (stat\_bin).

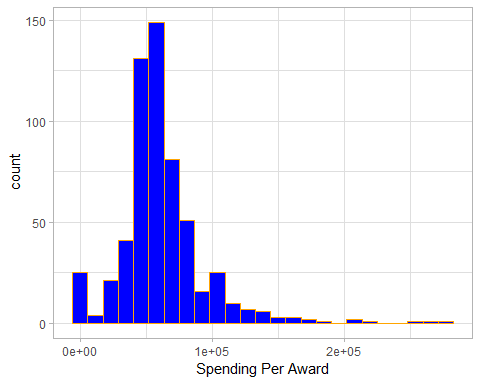
## Warning: Removed 2 rows containing missing values (geom\_bar).



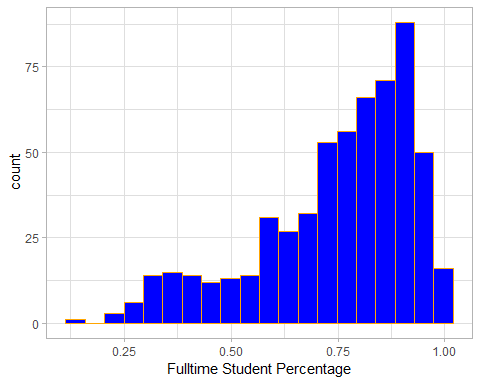
student <- ggplot(data = data, aes(x=student\_count)) +  
 geom\_histogram(bins = 20, fill = "blue", color = "orange")+  
 labs(x="Student Count")+  
 theme\_light()  
  
student



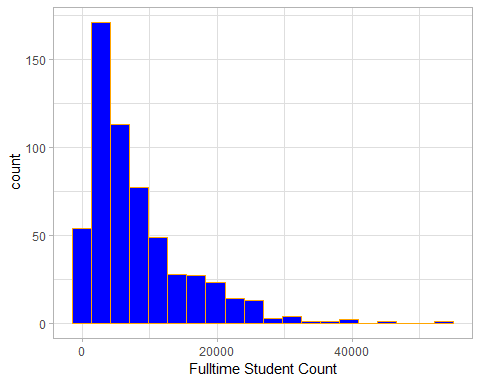
spending <- ggplot(data = data, aes(x=spending\_per\_award)) +  
 geom\_histogram(bins = 25, fill = "blue", color = "orange")+  
 labs(x="Spending Per Award")+  
 theme\_light()  
  
spending



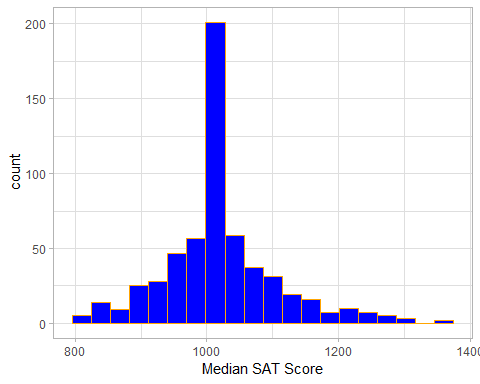
fullpct <- ggplot(data = data, aes(x=full\_time\_pct)) +  
 geom\_histogram(bins = 20, fill = "blue", color = "orange")+  
 labs(x="Fulltime Student Percentage")+  
 theme\_light()  
  
fullpct



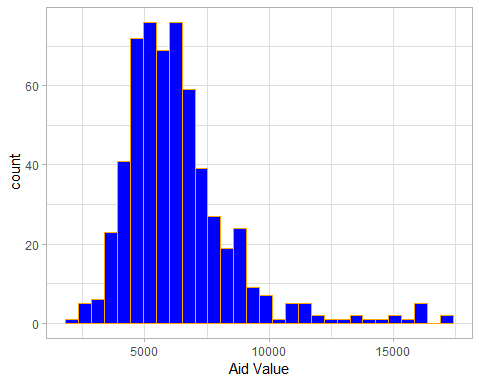
fullcount <- ggplot(data = data, aes(x=full\_time\_count)) +  
 geom\_histogram(bins = 20, fill = "blue", color = "orange")+  
 labs(x="Fulltime Student Count")+  
 theme\_light()  
  
fullcount



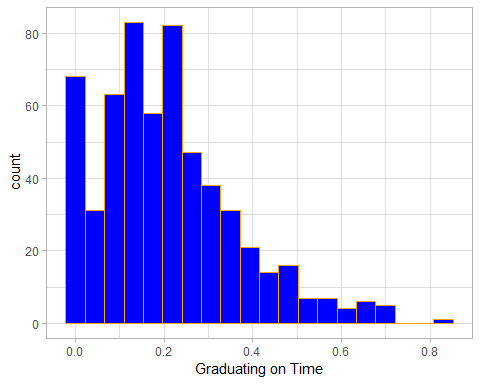
medsat <- ggplot(data = data, aes(x=med\_sat\_value)) +  
 geom\_histogram(bins = 20, fill = "blue", color = "orange")+  
 labs(x="Median SAT Score")+  
 theme\_light()  
  
medsat



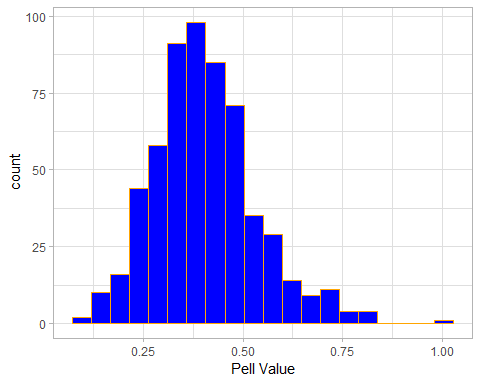
aid <- ggplot(data = data, aes(x=aid\_value)) +  
 geom\_histogram(bins = 30, fill = "blue", color = "orange")+  
 labs(x="Aid Value")+  
 theme\_light()  
  
aid



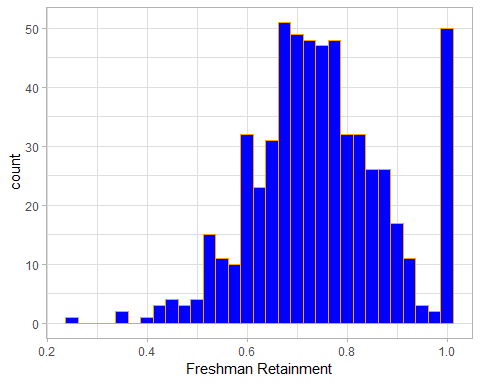
gradtime <- ggplot(data = data, aes(x=grad\_on\_time\_pct)) +  
 geom\_histogram(bins = 20, fill = "blue", color = "orange")+  
 labs(x="Graduating on Time")+  
 theme\_light()  
  
gradtime



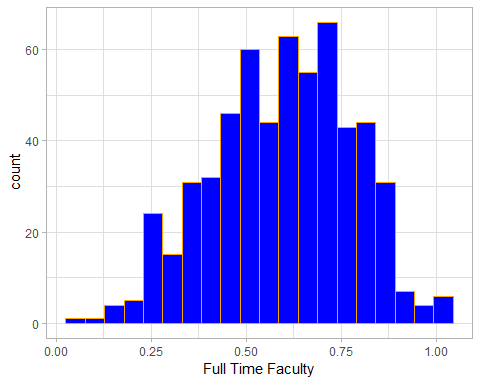
pell <- ggplot(data = data, aes(x=pell\_value)) +  
 geom\_histogram(bins = 20, fill = "blue", color = "orange")+  
 labs(x="Pell Value")+  
 theme\_light()  
  
pell



fresh <- ggplot(data = data, aes(x=fresh\_retain\_value)) +  
 geom\_histogram(binwidth = 0.025, fill = "blue", color = "orange")+  
 labs(x="Freshman Retainment")+  
 theme\_light()  
  
fresh

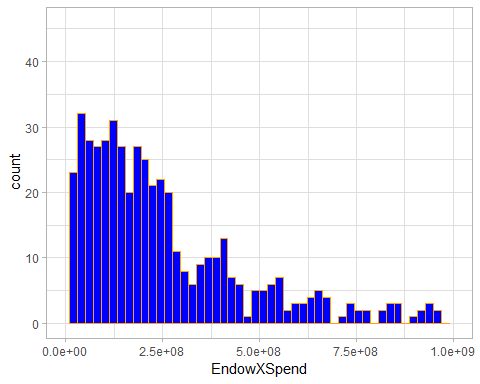


fullfac <- ggplot(data = data, aes(x=full\_time\_fac\_pct)) +  
 geom\_histogram(bins = 20, fill = "blue", color = "orange")+  
 labs(x="Full Time Faculty")+  
 theme\_light()  
  
fullfac

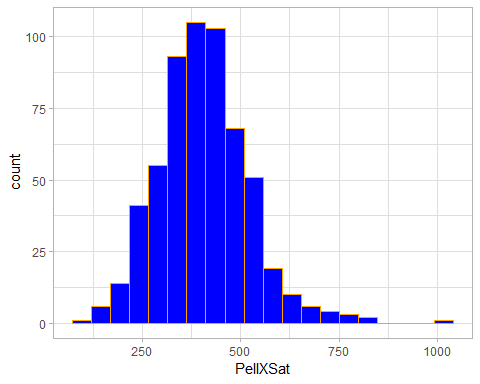


endowspend <- ggplot(data = data, aes(x=EndowXSpend)) +  
 geom\_histogram(bins = 50, fill = "blue", color = "orange")+  
 xlim(0,1000000000)+  
 theme\_light()  
  
endowspend

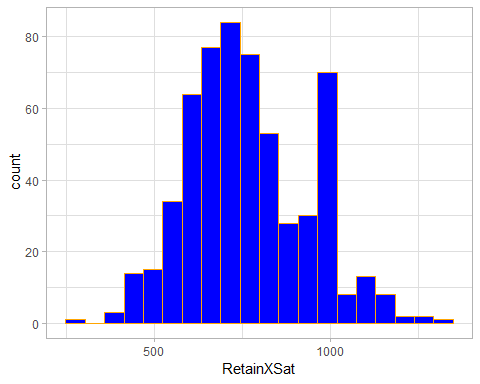
## Warning: Removed 56 rows containing non-finite values (stat\_bin).  
  
## Warning: Removed 2 rows containing missing values (geom\_bar).



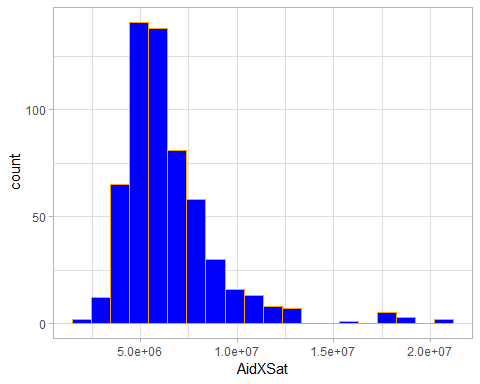
pellsat <- ggplot(data = data, aes(x=PellXSat)) +  
 geom\_histogram(bins = 20, fill = "blue", color = "orange")+  
 theme\_light()  
  
pellsat



retainsat <- ggplot(data = data, aes(x=RetainXSat)) +  
 geom\_histogram(bins = 20, fill = "blue", color = "orange")+  
 theme\_light()  
  
retainsat



aidsat <- ggplot(data = data, aes(x=AidXSat)) +  
 geom\_histogram(bins = 20, fill = "blue", color = "orange")+  
 theme\_light()  
  
aidsat



aidendow <- ggplot(data = data, aes(x=AidXEndow)) +  
 geom\_histogram(bins = 50, fill = "blue", color = "orange")+  
 xlim(0,100000000)+  
 theme\_light()  
  
aidendow

## Warning: Removed 52 rows containing non-finite values (stat\_bin).  
  
## Warning: Removed 2 rows containing missing values (geom\_bar).

