MC\_ENROLL

Juan Nunez

April 4, 2019

I set the working directory.

setwd("C:\\Users\\Juan Nunez\\Desktop\\DATA\_Capstone")

I open the packages.

library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

library(readr)  
library(ggplot2)  
library(tidyr)

I open the dataset.

mc\_enroll <- read\_csv("DATASETS\\Montgomery\_College\_Enrollment\_Data.csv")

## Parsed with column specification:  
## cols(  
## `Fall Term` = col\_double(),  
## `Student Type` = col\_character(),  
## `Student Status` = col\_character(),  
## Gender = col\_character(),  
## Ethnicity = col\_character(),  
## Race = col\_character(),  
## `Attending Germantown` = col\_character(),  
## `Attending Rockville` = col\_character(),  
## `Attending Takoma Park/SS` = col\_character(),  
## `Attend Day or Evening` = col\_character(),  
## `MC Program Description` = col\_character(),  
## `Age Group` = col\_character(),  
## `HS Category` = col\_character(),  
## `MCPS High School` = col\_character(),  
## `City in MD` = col\_character(),  
## State = col\_character(),  
## ZIP = col\_double(),  
## `County in MD` = col\_character()  
## )

Now let’s look at the top dataset

dim(mc\_enroll)

## [1] 25320 18

head(mc\_enroll)

## # A tibble: 6 x 18  
## `Fall Term` `Student Type` `Student Status` Gender Ethnicity Race   
## <dbl> <chr> <chr> <chr> <chr> <chr>  
## 1 2015 Continuing Full-Time Female Not Hisp~ White  
## 2 2015 Continuing Part-Time Male Not Hisp~ White  
## 3 2015 Continuing Part-Time Male Not Hisp~ Black  
## 4 2015 New Full-Time Male Not Hisp~ Asian  
## 5 2015 New Full-Time Female Hispanic White  
## 6 2015 Continuing Full-Time Female Hispanic Hisp~  
## # ... with 12 more variables: `Attending Germantown` <chr>, `Attending  
## # Rockville` <chr>, `Attending Takoma Park/SS` <chr>, `Attend Day or  
## # Evening` <chr>, `MC Program Description` <chr>, `Age Group` <chr>, `HS  
## # Category` <chr>, `MCPS High School` <chr>, `City in MD` <chr>,  
## # State <chr>, ZIP <dbl>, `County in MD` <chr>

Below is a code to see the entire dataset in terms of categories within variables.

lapply(mc\_enroll,table)

## $`Fall Term`  
##   
## 2015   
## 25320   
##   
## $`Student Type`  
##   
## Continuing HS Student New   
## 18513 485 4767   
## New/No HS Transfer Transfer no degree   
## 29 51 1132   
## Transfer with degree   
## 343   
##   
## $`Student Status`  
##   
## Full-Time Part-Time   
## 8890 16430   
##   
## $Gender  
##   
## Female Male Unknown   
## 13350 11963 7   
##   
## $Ethnicity  
##   
## Hispanic Not Hispanic Unknown   
## 6613 17984 723   
##   
## $Race  
##   
## Asian Black Hispanic Multi-Race   
## 3538 8217 2028 860   
## Native American Pacific Islander Unknown White   
## 501 293 52 9831   
##   
## $`Attending Germantown`  
##   
## No Yes   
## 18013 7307   
##   
## $`Attending Rockville`  
##   
## No Yes   
## 9034 16286   
##   
## $`Attending Takoma Park/SS`  
##   
## No Yes   
## 17445 7875   
##   
## $`Attend Day or Evening`  
##   
## Day & Evening Day Only Evening Only   
## 4867 16117 2737   
## No Specific Time[online]   
## 1599   
##   
## $`MC Program Description`  
##   
## Accounting (AA & AAS)   
## 345   
## Accounting (CT)   
## 108   
## Administrative Support Tech (CT)   
## 1   
## American Sign Language (AA & AAS)   
## 40   
## American Sign Language (CT)   
## 13   
## Applied Geography (AA & AAS)   
## 27   
## Architect. & Construct. Tech - Sustainability (LR)   
## 2   
## Architectural & Construction Tech (AA & AAS)   
## 230   
## Arts & Sciences Transfer (AA - All Tracks)   
## 1134   
## Arts & Sciences Transfer (CT)   
## 51   
## Automotive Technology (AA & AAS)   
## 197   
## Automotive Technology (CT)   
## 23   
## Biotechnology (AA & AAS)   
## 124   
## Biotechnology (CT)   
## 30   
## Building Trades Technology (AA & AAS)   
## 69   
## Building Trades Technology (CT)   
## 36   
## Building Trades Technology (LR)   
## 5   
## Business / International Business (AA)   
## 2289   
## Cartography & Geographic Ed / Info Sys (CT)   
## 11   
## Commun & Broadcasting Tech (AA & AAS - All Tracks)   
## 96   
## Commun & Broadcasting Tech (CT)   
## 121   
## Communication Studies (AA)   
## 249   
## Computer Applications (AA & AAS)   
## 139   
## Computer Applications (CT)   
## 65   
## Computer Gaming & Simulation (AA - All Tracks)   
## 252   
## Computer Graphics / Graphic Design (AAS)   
## 9   
## Computer Graphics / Graphic Design (CT)   
## 59   
## Computer Science - Computer Programming (CT)   
## 152   
## Computer Science & Technologies (AA - All Tracks)   
## 866   
## Credit (Undeclared / Undecided)   
## 1127   
## Criminal Justice (AA & AAS)   
## 638   
## Cybersecurity (AAS)   
## 297   
## Cybersecurity (CT)   
## 7   
## Diagnostic Medical Sonography (AA & AAS)   
## 50   
## Diagnostic Medical Sonography (CT)   
## 1   
## Digital Media & Web Technology (AAS)   
## 34   
## Digital Media & Web Technology (CT)   
## 41   
## Early Childhood Education (AA & AAS)   
## 104   
## Early Childhood Education (CT)   
## 161   
## Early Childhood Education (LR)   
## 3   
## Education / Teacher Education (AA & AAT)   
## 845   
## Electromechanical Sys Eng Tech (AA & AAS - Discnt)   
## 1   
## Eng Technologies (AA & AAS - Discontinued)   
## 1   
## Engineering Science (AA & AS - All Tracks)   
## 1482   
## Ethnic Social Studies (LR)   
## 1   
## Ethnic Studies (CT)   
## 3   
## Exercise Sci - Personal Trainer (CT)   
## 3   
## Exercise Sci - Personal Trainer (LR)   
## 12   
## Fire Sci./Preven., Emerg. Prepare. (AA, AS & AAS)   
## 27   
## Fire Sci./Preven., Emergency Prepare. (CT)   
## 5   
## Fire Science (LR)   
## 8   
## General Studies (AA - All Tracks)   
## 7449   
## Graphic Design (AA, AAS, & AFA - All Tracks)   
## 208   
## Graphic Design (AFA) - School of Art & Design   
## 22   
## Health Information Management (AA & AAS)   
## 74   
## Health Sciences (Pre-Clinical Studies)   
## 2969   
## Hospitality Management (AA & AAS)   
## 140   
## Hospitality Management (CT)   
## 40   
## Hospitality Management (LR)   
## 6   
## Information Systems Secirity   
## 3   
## Interior Design - PreProfessional (AAS)   
## 47   
## Interior Design (CT)   
## 26   
## Landscape Technology (AA & AAS)   
## 32   
## Landscape Technology (CT)   
## 23   
## Management (AA & AAS - All Tracks)   
## 3   
## Management (CT)   
## 33   
## Management (LR)   
## 5   
## Management of Construction (CT)   
## 30   
## Medical Coder/Abstractr/Biller (CT)   
## 1   
## Mental Health Associate (AA & AAS)   
## 101   
## Microcomputer Technician (AA & AAS)   
## 75   
## Microcomputer Technician (CT)   
## 5   
## Music Transfer (CT)   
## 24   
## Network & Wireless Technologies (CT)   
## 20   
## Network Engineer/Administration (CT)   
## 13   
## Nursing (AA & AAS)   
## 391   
## Paralegal Studies - Legal Analysis (LR)   
## 11   
## Paralegal Studies (AA & AAS)   
## 113   
## Paralegal Studies (CT)   
## 23   
## Photography (AA & AAS)   
## 77   
## Photography (CT)   
## 36   
## Physical Therapist Assistant (AAS)   
## 42   
## Polysomnography Technology (CT)   
## 16   
## Printing Management (AA & AAS)   
## 2   
## Printing Management (CT)   
## 1   
## Radiologic (X-Ray) Technology (AA & AAS)   
## 39   
## Recreation Leadership (AA)   
## 1   
## School of Art & Design - Applicants   
## 55   
## Science (AS - All Tracks)   
## 1432   
## Specialized Art Transfer (CT)   
## 17   
## Studio Art (AFA)   
## 87   
## Studio Art (AFA) - School of Art & Design   
## 10   
## Surgical Technologist (AAS)   
## 15   
## Technical Writing (CT)   
## 4   
## WIA (CE) Programs   
## 1   
## Women's Studies (CT)   
## 4   
##   
## $`Age Group`  
##   
## 20 or Younger 21 - 24 25 - 29 30 or Older Unknown   
## 10533 6349 3320 5116 2   
##   
## $`HS Category`  
##   
## Foreign Country GED Home-Schooled MCPS   
## 4910 516 229 13558   
## Mont.County Other Other Maryland Other State   
## 710 1614 3783   
##   
## $`MCPS High School`  
##   
## Albert Einstein HS & MC Art Cn Bethesda Chevy Chase High Schl   
## 655 357   
## Clarksburg High School Colonel Zadok Magruder HS   
## 576 688   
## Damascus High School Gaithersburg High School   
## 391 842   
## James Hubert Blake High School John F. Kennedy High School   
## 493 654   
## Montgomery Blair High School Northwest HS - Germantown   
## 852 739   
## Northwood High School Paint Branch High School   
## 478 553   
## Poolesville Jr-Sr High School Quince Orchard Sr High School   
## 197 526   
## Richard Montgomery High School Rockville High School   
## 633 501   
## Seneca Valley High School Sherwood High School   
## 545 602   
## Springbrook Sr High School Thomas Sprigg Wootton High Sch   
## 613 442   
## Walt Whitman High School Walter Johnson High School   
## 231 596   
## Watkins Mill High School Wheaton High School   
## 602 483   
## Winston Churchill High School   
## 309   
##   
## $`City in MD`  
##   
## Abingdon Accokeek Adamstown   
## 1 1 2   
## Adelphi Albany Aldie   
## 57 1 1   
## Alexandria Alva Annandale   
## 6 1 2   
## Annapolis Arlington Ashburn   
## 9 15 4   
## Ashton Aspen Hill Atlanta   
## 27 1 1   
## Baltimore Barnesville Beallsville   
## 16 7 3   
## Bellefonte Beltsville Berlin   
## 1 121 1   
## Berwyn Heights Betheada Bethesda   
## 4 1 812   
## Bladensburg Boca Raton Bolling AFB   
## 18 1 1   
## Boonsboro Bowie Boyds   
## 2 64 223   
## Brandywine Brentwood Bridgeville   
## 1 20 1   
## Brinklow Bronx Brookeville   
## 6 1 149   
## Brooklyn Brunswick Burke   
## 2 1 1   
## Burtonsville Cabin John California   
## 305 18 1   
## Camillus Camp Springs Capitol Heights   
## 1 2 28   
## Carrboro Catonsville Centreville   
## 1 1 4   
## Chantilly Chapel Hill Charleston   
## 1 1 1   
## Chesapeake Beach Chestertown Cheverly   
## 1 1 10   
## Chevy Chase Chillum Clarkburg   
## 193 1 1   
## Clarksbrug Clarksburg Clear Spring   
## 1 429 1   
## Clifton Clinton Cockeysville   
## 1 15 1   
## College Park Colonial Heights Columbia   
## 71 1 19   
## Columbus Cooksville Dallas   
## 1 2 1   
## Damascus Darnestown Derwood   
## 355 12 336   
## Dickerson District Heights Dover   
## 35 22 1   
## Dunkirk Dunn Loring East Liverpool   
## 1 1 1   
## Easton Edgewood Elkridge   
## 2 4 3   
## Ellicott City Fairfax Fairmount Heights   
## 7 1 2   
## Falls Church Farmville Fayetteville   
## 2 1 2   
## Fort Washington Frederick Fulton   
## 24 84 1   
## Gaitersburg Gaitherburg Gaithersburg   
## 1 2 3220   
## Gaithursburg Garrett Park Garrett Parkd   
## 1 11 1   
## Germantown Germatown Germnatown   
## 2675 2 1   
## Gettysburg Glen Burnie Glen Echo   
## 1 5 2   
## Glenarden Glenn Dale Glenwood   
## 7 11 1   
## Graham Grand Rapids Grasonville   
## 1 2 2   
## Greenbelt Greenwich Hagerstown   
## 58 1 14   
## Hancock Hanover Harpers Ferry   
## 1 2 1   
## Harrisburg Harrisonburg Hercules   
## 1 1 1   
## Herndon Highland Howard City   
## 2 5 1   
## Hughesville Hyattsville Ijamsville   
## 1 327 11   
## Indian Head Irmo Irving   
## 2 1 1   
## Jessup Kensington Knoxville   
## 1 312 1   
## Landover Lanham Larbo   
## 5 67 1   
## Largo Laurel Laytonsville   
## 1 115 48   
## Lovettsville Lutherville Timonium Manassas Park   
## 1 1 1   
## Marbury Marietta Marriottsville   
## 1 1 1   
## Marshfield Martinsburg Mc Lean   
## 1 4 1   
## McLean Merrick Middle River   
## 1 1 1   
## Middletown Millersville Millsboro   
## 6 2 1   
## Mitchellville Monrovia Montgomery Village   
## 1 5 901   
## Mount Airy Mount Rainier Myersville   
## 26 33 3   
## Nairobi Nashville New Carrollton   
## 1 1 32   
## New Hope New Market New York   
## 1 9 1   
## Newark North Bethesda North Brentwood   
## 1 9 1   
## North Potomac Nottingham Odenton   
## 343 2 1   
## Olney OLNEY Owings Mills   
## 560 1 4   
## Oxon Hill Parkville Pasadena   
## 16 3 1   
## Philadelphia Pikesville Pittsburgh   
## 1 2 2   
## Poolesville Potomac Prince Frederick   
## 145 550 2   
## Quantico Reisterstown Riverdale   
## 1 3 43   
## Rockille Rockville Rohrersville   
## 1 2925 1   
## Rosedale Roswell Saint Charles   
## 1 1 1   
## Salisbury San Fernando San Francisco   
## 2 1 1   
## San Juan Sandy Spring Savannah   
## 1 39 1   
## Severn Shepherdstown SIiver Spring   
## 2 1 1   
## Silver Sping silver spring Silver Spring   
## 1 1 7464   
## Sliver Spring Smyrna Spencerville   
## 1 2 12   
## Springdale Springfield Sterling   
## 2 3 2   
## Suitland Sykesville Syracuse   
## 19 6 1   
## Takoma Park Tampa Temple Hills   
## 586 1 22   
## The Plains Towson Twinsburg   
## 1 3 1   
## University Park Upper Malboro Upper Marlboro   
## 6 2 54   
## Van Nuys Ventura Vienna   
## 1 1 8   
## Wagener Waldorf Walkersville   
## 1 11 1   
## Walnut Creek Washington Washington Grove   
## 1 726 16   
## Waynesboro West Friendship Westminster   
## 1 1 4   
## Westport Wheaton White Plains   
## 1 7 2   
## Williamsport Winchester Windsor   
## 1 1 1   
## Windsor Mill Woodbine Woodbridge   
## 5 4 5   
## Woodsboro   
## 1   
##   
## $State  
##   
## CA CT DC DE FL GA MA MD MI MO NC NJ   
## 6 2 727 3 2 6 1 24453 3 2 5 1   
## NY OH OK PA PR SC TN TX VA WV   
## 7 3 1 9 1 3 1 3 67 6   
##   
## $ZIP  
##   
## 926 2050 6830 6880 7102 10028 10463 11422 11566 12203 13031 13215   
## 1 1 1 1 1 1 1 1 1 1 1 1   
## 15217 15235 16823 17102 17222 17268 17325 18938 19129 19904 19933 19966   
## 1 1 1 1 1 1 1 1 1 1 1 1   
## 20001 20002 20003 20004 20005 20007 20008 20009 20010 20011 20012 20015   
## 41 46 19 3 6 20 23 52 57 179 52 23   
## 20016 20017 20018 20019 20020 20024 20032 20036 20037 20039 20040 20105   
## 37 34 34 38 33 9 14 1 4 1 1 1   
## 20111 20120 20121 20124 20147 20148 20152 20165 20166 20170 20171 20180   
## 1 2 2 1 2 2 1 1 1 1 1 1   
## 20198 20601 20602 20603 20607 20613 20619 20637 20640 20658 20678 20695   
## 1 2 2 7 1 1 1 1 2 1 2 2   
## 20705 20706 20707 20708 20710 20712 20715 20716 20720 20721 20722 20723   
## 121 73 55 37 18 34 5 11 21 28 21 10   
## 20724 20732 20734 20735 20737 20740 20741 20743 20744 20745 20746 20747   
## 12 1 1 14 42 74 1 30 24 16 19 22   
## 20748 20754 20759 20769 20770 20772 20774 20777 20781 20782 20783 20784   
## 24 1 1 11 59 27 35 5 24 125 182 62   
## 20785 20787 20794 20810 20812 20814 20815 20816 20817 20818 20830 20832   
## 45 1 1 1 2 259 194 103 436 18 2 557   
## 20833 20835 20837 20838 20839 20841 20842 20845 20847 20848 20849 20850   
## 149 1 145 7 3 223 34 1 1 3 5 991   
## 20851 20852 20853 20854 20855 20860 20861 20862 20866 20868 20871 20872   
## 427 726 758 550 358 39 27 6 305 12 429 354   
## 20874 20875 20876 20877 20878 20879 20880 20882 20883 20886 20889 20895   
## 1828 7 847 1056 1283 873 16 334 1 975 16 311   
## 20896 20898 20901 20902 20903 20904 20905 20906 20907 20909 20910 20912   
## 12 2 888 1309 720 1611 415 1738 2 1 782 584   
## 20914 20915 20916 20918 20978 21009 21030 21040 21042 21043 21044 21045   
## 4 3 1 4 1 1 1 4 4 3 6 9   
## 21046 21060 21061 21075 21076 21093 21104 21108 21113 21117 21122 21136   
## 3 2 3 3 2 1 1 2 1 4 1 3   
## 21144 21157 21158 21201 21204 21206 21208 21211 21212 21215 21217 21218   
## 2 3 1 1 2 2 2 1 1 1 1 1   
## 21220 21223 21224 21225 21228 21229 21234 21236 21239 21244 21286 21401   
## 1 1 1 2 1 3 3 2 3 5 1 1   
## 21403 21404 21409 21601 21620 21638 21701 21702 21703 21704 21710 21713   
## 2 2 4 2 1 2 13 17 25 29 2 2   
## 21716 21722 21723 21738 21740 21742 21750 21754 21758 21769 21770 21771   
## 1 1 2 1 9 5 1 11 1 6 5 26   
## 21773 21774 21779 21784 21793 21794 21795 21797 21798 21801 21811 21856   
## 3 9 1 6 1 1 1 4 1 2 1 1   
## 22003 22015 22027 22033 22041 22046 22102 22152 22153 22181 22182 22191   
## 2 1 1 1 1 1 2 2 1 2 6 1   
## 22192 22193 22201 22202 22203 22204 22205 22206 22209 22304 22307 22309   
## 2 2 3 2 4 2 1 1 1 2 1 1   
## 22314 22315 22601 22801 23834 23901 25403 25425 25443 27253 27510 27517   
## 1 1 1 1 1 1 4 1 1 1 1 1   
## 28311 28328 29063 29164 29406 30062 30076 30080 30342 31405 33496 33614   
## 1 1 1 1 1 1 1 2 1 1 1 1   
## 37206 43235 43920 44087 49329 49504 49506 63301 65203 73717 75038 75230   
## 1 1 1 1 1 1 1 1 1 1 1 1   
## 76015 91405 93003 94116 94547 94597 95492   
## 1 1 1 1 1 1 1   
##   
## $`County in MD`  
##   
## Anne Arundel Baltimore Baltimore City Calvert   
## 36 26 16 4   
## Carroll Charles Frederick Harford   
## 15 17 150 5   
## Howard Kent Montgomery Other   
## 49 1 22840 801   
## Prince George's Queen Anne's St. Mary's Talbot   
## 1263 2 1 2   
## Washington Wicomico Worcester   
## 88 3 1

Below I add an underscore to the names of the columns.

mc\_enroll2 <- mc\_enroll  
colnames(mc\_enroll2) = gsub(" ", "\_",colnames(mc\_enroll2))

I create a dummy for whether they are continuing or new to MC

mc\_enroll2['MC\_TRUE']= "Yes"  
mc\_enroll2$MC\_TRUE[mc\_enroll2$Student\_Type != "Continuing"] <- "No"  
table(mc\_enroll2$MC\_TRUE)

##   
## No Yes   
## 6807 18513

Take out the unknowns from Race Ethnicity and Gender

mc\_enroll2$HISPANIC <- mc\_enroll2$Ethnicity  
mc\_enroll2$HISPANIC[mc\_enroll2$Ethnicity == "Unknown"] <- NA  
table(mc\_enroll2$HISPANIC)

##   
## Hispanic Not Hispanic   
## 6613 17984

mc\_enroll2$RACE <- mc\_enroll2$Race  
mc\_enroll2$RACE[mc\_enroll2$Race == "Unknown"] <- NA  
table(mc\_enroll2$RACE)

##   
## Asian Black Hispanic Multi-Race   
## 3538 8217 2028 860   
## Native American Pacific Islander White   
## 501 293 9831

mc\_enroll2['SEX']= "F"  
mc\_enroll2$SEX[mc\_enroll2$Gender == "Male"] <- "M"  
mc\_enroll2$SEX[mc\_enroll2$Gender == "Unknown"] <- NA  
table(mc\_enroll2$SEX)

##   
## F M   
## 13350 11963

Below I create a 3 category variable for MC campus.

mc\_enroll2['MC\_CAMPUS']= "Other"  
mc\_enroll2$MC\_CAMPUS[mc\_enroll2$`Attending\_Takoma\_Park/SS` == "Yes"] <- "TP - SS"  
mc\_enroll2$MC\_CAMPUS[mc\_enroll2$Attending\_Rockville == "Yes"] <- "RV"  
mc\_enroll2$MC\_CAMPUS[mc\_enroll2$Attending\_Germantown == "Yes"] <- "GT"  
table(mc\_enroll2$MC\_CAMPUS)

##   
## GT RV TP - SS   
## 7307 13110 4903

Below I make unknown NA in the AGE variable

mc\_enroll2$AGE\_GROUP\_fixed <- mc\_enroll2$Age\_Group  
mc\_enroll2$AGE\_GROUP\_fixed[mc\_enroll2$AGE\_GROUP\_fixed == "Unknown"] <- NA  
table(mc\_enroll2$AGE\_GROUP\_fixed)

##   
## 20 or Younger 21 - 24 25 - 29 30 or Older   
## 10533 6349 3320 5116

Below I create a dummy for MCPS.

mc\_enroll2['MCPS\_DUM']= "Other"  
mc\_enroll2$MCPS\_DUM[mc\_enroll2$HS\_Category == "MCPS"] <- "MCPS"  
table(mc\_enroll2$MCPS\_DUM)

##   
## MCPS Other   
## 13558 11762

Below I create a dummy for State

mc\_enroll2['STATE\_DUM']= "Other"  
mc\_enroll2$STATE\_DUM[mc\_enroll2$State == "MD"] <- "MD"  
table(mc\_enroll2$STATE\_DUM)

##   
## MD Other   
## 24453 867

I fix the zip codes.

class(mc\_enroll2$ZIP)

## [1] "numeric"

mc\_enroll2$ZIP <- as.character(mc\_enroll2$ZIP)  
class(mc\_enroll2$ZIP)

## [1] "character"

table(mc\_enroll2$ZIP)

##   
## 10028 10463 11422 11566 12203 13031 13215 15217 15235 16823 17102 17222   
## 1 1 1 1 1 1 1 1 1 1 1 1   
## 17268 17325 18938 19129 19904 19933 19966 20001 20002 20003 20004 20005   
## 1 1 1 1 1 1 1 41 46 19 3 6   
## 20007 20008 20009 20010 20011 20012 20015 20016 20017 20018 20019 20020   
## 20 23 52 57 179 52 23 37 34 34 38 33   
## 20024 20032 20036 20037 20039 20040 20105 20111 20120 20121 20124 20147   
## 9 14 1 4 1 1 1 1 2 2 1 2   
## 20148 20152 20165 20166 20170 20171 20180 20198 2050 20601 20602 20603   
## 2 1 1 1 1 1 1 1 1 2 2 7   
## 20607 20613 20619 20637 20640 20658 20678 20695 20705 20706 20707 20708   
## 1 1 1 1 2 1 2 2 121 73 55 37   
## 20710 20712 20715 20716 20720 20721 20722 20723 20724 20732 20734 20735   
## 18 34 5 11 21 28 21 10 12 1 1 14   
## 20737 20740 20741 20743 20744 20745 20746 20747 20748 20754 20759 20769   
## 42 74 1 30 24 16 19 22 24 1 1 11   
## 20770 20772 20774 20777 20781 20782 20783 20784 20785 20787 20794 20810   
## 59 27 35 5 24 125 182 62 45 1 1 1   
## 20812 20814 20815 20816 20817 20818 20830 20832 20833 20835 20837 20838   
## 2 259 194 103 436 18 2 557 149 1 145 7   
## 20839 20841 20842 20845 20847 20848 20849 20850 20851 20852 20853 20854   
## 3 223 34 1 1 3 5 991 427 726 758 550   
## 20855 20860 20861 20862 20866 20868 20871 20872 20874 20875 20876 20877   
## 358 39 27 6 305 12 429 354 1828 7 847 1056   
## 20878 20879 20880 20882 20883 20886 20889 20895 20896 20898 20901 20902   
## 1283 873 16 334 1 975 16 311 12 2 888 1309   
## 20903 20904 20905 20906 20907 20909 20910 20912 20914 20915 20916 20918   
## 720 1611 415 1738 2 1 782 584 4 3 1 4   
## 20978 21009 21030 21040 21042 21043 21044 21045 21046 21060 21061 21075   
## 1 1 1 4 4 3 6 9 3 2 3 3   
## 21076 21093 21104 21108 21113 21117 21122 21136 21144 21157 21158 21201   
## 2 1 1 2 1 4 1 3 2 3 1 1   
## 21204 21206 21208 21211 21212 21215 21217 21218 21220 21223 21224 21225   
## 2 2 2 1 1 1 1 1 1 1 1 2   
## 21228 21229 21234 21236 21239 21244 21286 21401 21403 21404 21409 21601   
## 1 3 3 2 3 5 1 1 2 2 4 2   
## 21620 21638 21701 21702 21703 21704 21710 21713 21716 21722 21723 21738   
## 1 2 13 17 25 29 2 2 1 1 2 1   
## 21740 21742 21750 21754 21758 21769 21770 21771 21773 21774 21779 21784   
## 9 5 1 11 1 6 5 26 3 9 1 6   
## 21793 21794 21795 21797 21798 21801 21811 21856 22003 22015 22027 22033   
## 1 1 1 4 1 2 1 1 2 1 1 1   
## 22041 22046 22102 22152 22153 22181 22182 22191 22192 22193 22201 22202   
## 1 1 2 2 1 2 6 1 2 2 3 2   
## 22203 22204 22205 22206 22209 22304 22307 22309 22314 22315 22601 22801   
## 4 2 1 1 1 2 1 1 1 1 1 1   
## 23834 23901 25403 25425 25443 27253 27510 27517 28311 28328 29063 29164   
## 1 1 4 1 1 1 1 1 1 1 1 1   
## 29406 30062 30076 30080 30342 31405 33496 33614 37206 43235 43920 44087   
## 1 1 1 2 1 1 1 1 1 1 1 1   
## 49329 49504 49506 63301 65203 6830 6880 7102 73717 75038 75230 76015   
## 1 1 1 1 1 1 1 1 1 1 1 1   
## 91405 926 93003 94116 94547 94597 95492   
## 1 1 1 1 1 1 1

mc\_enroll2$ZIP[mc\_enroll2$ZIP == "926"] <- "00926"  
mc\_enroll2$ZIP[mc\_enroll2$ZIP == "2050"] <- "02050"  
mc\_enroll2$ZIP[mc\_enroll2$ZIP == "6830"] <- "06830"  
mc\_enroll2$ZIP[mc\_enroll2$ZIP == "7102"] <- "07102"  
table(mc\_enroll2$ZIP)

##   
## 00926 02050 06830 07102 10028 10463 11422 11566 12203 13031 13215 15217   
## 1 1 1 1 1 1 1 1 1 1 1 1   
## 15235 16823 17102 17222 17268 17325 18938 19129 19904 19933 19966 20001   
## 1 1 1 1 1 1 1 1 1 1 1 41   
## 20002 20003 20004 20005 20007 20008 20009 20010 20011 20012 20015 20016   
## 46 19 3 6 20 23 52 57 179 52 23 37   
## 20017 20018 20019 20020 20024 20032 20036 20037 20039 20040 20105 20111   
## 34 34 38 33 9 14 1 4 1 1 1 1   
## 20120 20121 20124 20147 20148 20152 20165 20166 20170 20171 20180 20198   
## 2 2 1 2 2 1 1 1 1 1 1 1   
## 20601 20602 20603 20607 20613 20619 20637 20640 20658 20678 20695 20705   
## 2 2 7 1 1 1 1 2 1 2 2 121   
## 20706 20707 20708 20710 20712 20715 20716 20720 20721 20722 20723 20724   
## 73 55 37 18 34 5 11 21 28 21 10 12   
## 20732 20734 20735 20737 20740 20741 20743 20744 20745 20746 20747 20748   
## 1 1 14 42 74 1 30 24 16 19 22 24   
## 20754 20759 20769 20770 20772 20774 20777 20781 20782 20783 20784 20785   
## 1 1 11 59 27 35 5 24 125 182 62 45   
## 20787 20794 20810 20812 20814 20815 20816 20817 20818 20830 20832 20833   
## 1 1 1 2 259 194 103 436 18 2 557 149   
## 20835 20837 20838 20839 20841 20842 20845 20847 20848 20849 20850 20851   
## 1 145 7 3 223 34 1 1 3 5 991 427   
## 20852 20853 20854 20855 20860 20861 20862 20866 20868 20871 20872 20874   
## 726 758 550 358 39 27 6 305 12 429 354 1828   
## 20875 20876 20877 20878 20879 20880 20882 20883 20886 20889 20895 20896   
## 7 847 1056 1283 873 16 334 1 975 16 311 12   
## 20898 20901 20902 20903 20904 20905 20906 20907 20909 20910 20912 20914   
## 2 888 1309 720 1611 415 1738 2 1 782 584 4   
## 20915 20916 20918 20978 21009 21030 21040 21042 21043 21044 21045 21046   
## 3 1 4 1 1 1 4 4 3 6 9 3   
## 21060 21061 21075 21076 21093 21104 21108 21113 21117 21122 21136 21144   
## 2 3 3 2 1 1 2 1 4 1 3 2   
## 21157 21158 21201 21204 21206 21208 21211 21212 21215 21217 21218 21220   
## 3 1 1 2 2 2 1 1 1 1 1 1   
## 21223 21224 21225 21228 21229 21234 21236 21239 21244 21286 21401 21403   
## 1 1 2 1 3 3 2 3 5 1 1 2   
## 21404 21409 21601 21620 21638 21701 21702 21703 21704 21710 21713 21716   
## 2 4 2 1 2 13 17 25 29 2 2 1   
## 21722 21723 21738 21740 21742 21750 21754 21758 21769 21770 21771 21773   
## 1 2 1 9 5 1 11 1 6 5 26 3   
## 21774 21779 21784 21793 21794 21795 21797 21798 21801 21811 21856 22003   
## 9 1 6 1 1 1 4 1 2 1 1 2   
## 22015 22027 22033 22041 22046 22102 22152 22153 22181 22182 22191 22192   
## 1 1 1 1 1 2 2 1 2 6 1 2   
## 22193 22201 22202 22203 22204 22205 22206 22209 22304 22307 22309 22314   
## 2 3 2 4 2 1 1 1 2 1 1 1   
## 22315 22601 22801 23834 23901 25403 25425 25443 27253 27510 27517 28311   
## 1 1 1 1 1 4 1 1 1 1 1 1   
## 28328 29063 29164 29406 30062 30076 30080 30342 31405 33496 33614 37206   
## 1 1 1 1 1 1 2 1 1 1 1 1   
## 43235 43920 44087 49329 49504 49506 63301 65203 6880 73717 75038 75230   
## 1 1 1 1 1 1 1 1 1 1 1 1   
## 76015 91405 93003 94116 94547 94597 95492   
## 1 1 1 1 1 1 1

Clean County in MD.

mc\_enroll2['MD\_COUNTY']= "Other"  
mc\_enroll2$MD\_COUNTY[mc\_enroll2$County\_in\_MD == "Montgomery"] <- "Montgomery"  
head(mc\_enroll2)

## # A tibble: 6 x 27  
## Fall\_Term Student\_Type Student\_Status Gender Ethnicity Race   
## <dbl> <chr> <chr> <chr> <chr> <chr>  
## 1 2015 Continuing Full-Time Female Not Hisp~ White  
## 2 2015 Continuing Part-Time Male Not Hisp~ White  
## 3 2015 Continuing Part-Time Male Not Hisp~ Black  
## 4 2015 New Full-Time Male Not Hisp~ Asian  
## 5 2015 New Full-Time Female Hispanic White  
## 6 2015 Continuing Full-Time Female Hispanic Hisp~  
## # ... with 21 more variables: Attending\_Germantown <chr>,  
## # Attending\_Rockville <chr>, `Attending\_Takoma\_Park/SS` <chr>,  
## # Attend\_Day\_or\_Evening <chr>, MC\_Program\_Description <chr>,  
## # Age\_Group <chr>, HS\_Category <chr>, MCPS\_High\_School <chr>,  
## # City\_in\_MD <chr>, State <chr>, ZIP <chr>, County\_in\_MD <chr>,  
## # MC\_TRUE <chr>, HISPANIC <chr>, RACE <chr>, SEX <chr>, MC\_CAMPUS <chr>,  
## # AGE\_GROUP\_fixed <chr>, MCPS\_DUM <chr>, STATE\_DUM <chr>,  
## # MD\_COUNTY <chr>

table(mc\_enroll2$MD\_COUNTY)

##   
## Montgomery Other   
## 22840 2480

Below I save teh dataset to csv.

write.csv(mc\_enroll2, file = "mc\_enroll2.csv",row.names=FALSE, na="")

table(mc\_enroll2$Student\_Type)

##   
## Continuing HS Student New   
## 18513 485 4767   
## New/No HS Transfer Transfer no degree   
## 29 51 1132   
## Transfer with degree   
## 343

table(mc\_enroll2$Student\_Status)

##   
## Full-Time Part-Time   
## 8890 16430

table(mc\_enroll2$RACE)

##   
## Asian Black Hispanic Multi-Race   
## 3538 8217 2028 860   
## Native American Pacific Islander White   
## 501 293 9831

table(mc\_enroll2$Race)

##   
## Asian Black Hispanic Multi-Race   
## 3538 8217 2028 860   
## Native American Pacific Islander Unknown White   
## 501 293 52 9831

mc\_enroll2$RACE <- mc\_enroll2$Race  
  
mc\_enroll2$RACE[mc\_enroll2$RACE == "Multi-Race"] <- "Other"  
mc\_enroll2$RACE[mc\_enroll2$RACE == "Native American"] <- "Other"  
mc\_enroll2$RACE[mc\_enroll2$RACE == "Pacific Islander"] <- "Other"  
mc\_enroll2$RACE[mc\_enroll2$RACE == "Unknown"] <- "Other"  
  
table(mc\_enroll2$RACE)

##   
## Asian Black Hispanic Other White   
## 3538 8217 2028 1706 9831

table(mc\_enroll2$AGE\_GROUP\_fixed)

##   
## 20 or Younger 21 - 24 25 - 29 30 or Older   
## 10533 6349 3320 5116

mc\_enroll2$RACE\_n <- factor(mc\_enroll2$RACE, levels = c("White", "Black", "Hispanic", "Asian", "Other"))  
table(mc\_enroll2$RACE\_n)

##   
## White Black Hispanic Asian Other   
## 9831 8217 2028 3538 1706

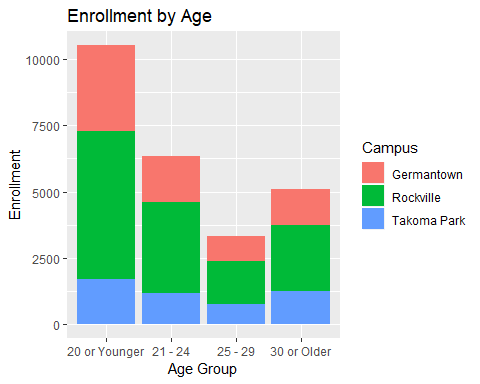
table(mc\_enroll2$RACE)

##   
## Asian Black Hispanic Other White   
## 3538 8217 2028 1706 9831

table(mc\_enroll2$MC\_CAMPUS)

##   
## GT RV TP - SS   
## 7307 13110 4903

mc\_enroll2 %>% filter(!is.na(AGE\_GROUP\_fixed)) %>%  
ggplot(aes(x = AGE\_GROUP\_fixed, fill = MC\_CAMPUS)) +  
geom\_bar() +  
labs(title = "Enrollment by Age", x= "Age Group", y = "Enrollment") +  
guides(fill=guide\_legend(title = "Campus")) +  
scale\_fill\_discrete(labels = c("Germantown", "Rockville", "Takoma Park"))



mc\_enroll2 %>% filter(!is.na(MC\_CAMPUS)) %>%  
ggplot() +  
aes(x = MC\_CAMPUS , fill = Student\_Status) +  
geom\_bar(position = 'dodge') +  
scale\_fill\_discrete(labels = c("Full-Time", "Part-Time")) +  
labs(fill = "Status", y = "Enrollment", x ="Campus", title = "Enrollment by Campus and Status") +  
scale\_x\_discrete(labels = c("Germantown" , "Rockville", "Takoma Park"))

