IE 48A - Final

Ecemnaz Bay

9/12/2020

Part I: Short and Simple Questions

1.Comments on COVID Analyses

It is so hard to conduct analysis in such an unknown and changing situation. Making comparisons between countries is even harder as population, life style and data quality is changing widely between different countries. Analyzing only number of confirmed cases or number of deaths will be misleading. Including the population size is a must to have a meaningful outcome. In that manner, cases can be compared per million population. Moreover, to have some idea about quality of medical system and quality of the data, country welfare can be included in the studies. Another important points is day to day changing situations. Before coming up with an output, analysis should indicate a trend. Analysts may consider to select countries for analysis after a specific minimum number of days since the first case like 3 weeks. Considering a minimum number of cases to select a country for analysis can also be helpful to include countries with some sense of past trends. Testing strategies of countries have also a on the numbers. In order to take that impact into account, analysis should include same rates with test numbers (like number of cases or deaths per test). In a perfect word that we can reach out the real number of cases and death, it would be very useful to analyse the fatality rate. As that is highly biased and dependent to the country policies, an estimation model can be used to have insight. Estimation of the expected number of deaths in analyzed countries at a given period of time may help to indicate the unexpected (most probably because of pandemic) increase of deaths. That approach would be more sophisticated and complex, yet it would decrease the bias.

2.Exploratory data analysis workflow

I start exploration with examining data structure, print number of rows and columns together with their data types. According to data types I would choose the useful visualization ways. For instance, I can plot trendlines to detect seasonality in time series data or heatmaps to see correlations. I can basically find out if data has some odd patterns or characteristics. Moreover, I use scatter diagrams and boxplots to detect outliers. After outlier detection, I plot the densities or histograms to understand the underlying distributions and patterns behind the data, if there are any.

That visualizations feeds the data cleaning process as well. I also consider if there are duplications or null values. I singularize the data and remove or find an imputation method (more preferably) for the null values. The abundance of null records will cause gaps in the analysis, so I try to be more careful before getting insights from the columns having more missing points. After that cleaning phase I go back to visualizations to see if I reached to a more meaningful data set and start to the analysis.

Afterwards, I would try to extract the most impactful issue that society is having trouble with. I think, that is how I can provide social benefit. For example, in societies where people die of malnutrition or diarrhea, social benefit of investment in gender equality or education will not be very high. With remaining amount, I would consider investing in subjects that interact with each other. If I got positive correlations between two subjects, I certainly invest one of them. For example, if I see that investment in education will affect gender equality or investment in employment prevents poverty in long terms, I allocate more money for these areas. The most basic metrics such as death rate, poverty, literacy rate would work well to examine the impact. I prefer to be data-oriented rather than moving on with personal perceptions while doing budget allocation.

3.Star Wars Data Visualization

3.a Getting a Glimpse of the Data

```
starwars %>% unnest(films) %>% unnest(vehicles) %>% unnest(starships) %>% glimpse()
```

```
## Rows: 83
## Columns: 14
## $ name       <chr> "Luke Skywalker", "Luke Skywalker", "Luke Skywalker", "L...
## $ height     <int> 172, 172, 172, 172, 172, 172, 172, 172, 172, 172, 172, 1...
## $ mass       <dbl> 77, 77, 77, 77, 77, 77, 77, 77, 77, 77, 77, 77, 77, 77, ...
## $ hair_color <chr> "blond", "blond", "blond", "blond", "blond", "blond", "b...
## $ skin_color <chr> "fair", "fair", "fair", "fair", "fair", "fair", "fair", ...
## $ eye_color  <chr> "blue", "blue", "blue", "blue", "blue", "blue", "blue", ...
## $ birth_year <dbl> 19, 19, 19, 19, 19, 19, 19, 19, 19, 19, 19, 19, 19, 19, ...
## $ sex        <chr> "male", "male", "male", "male", "male", "male", "male", ...
## $ gender     <chr> "masculine", "masculine", "masculine", "masculine", "mas...
## $ homeworld  <chr> "Tatooine", "Tatooine", "Tatooine", "Tatooine", "Tatooin...
## $ species    <chr> "Human", "Human", "Human", "Human", "Human", "Human", "H...
## $ films      <chr> "The Empire Strikes Back", "The Empire Strikes Back", "T...
## $ vehicles   <chr> "Snowspeeder", "Snowspeeder", "Imperial Speeder Bike", "...
## $ starships  <chr> "X-wing", "Imperial shuttle", "X-wing", "Imperial shuttl...
```

3.b Star Wars Characters Analysis

I wanted to see 3 main things from the data :

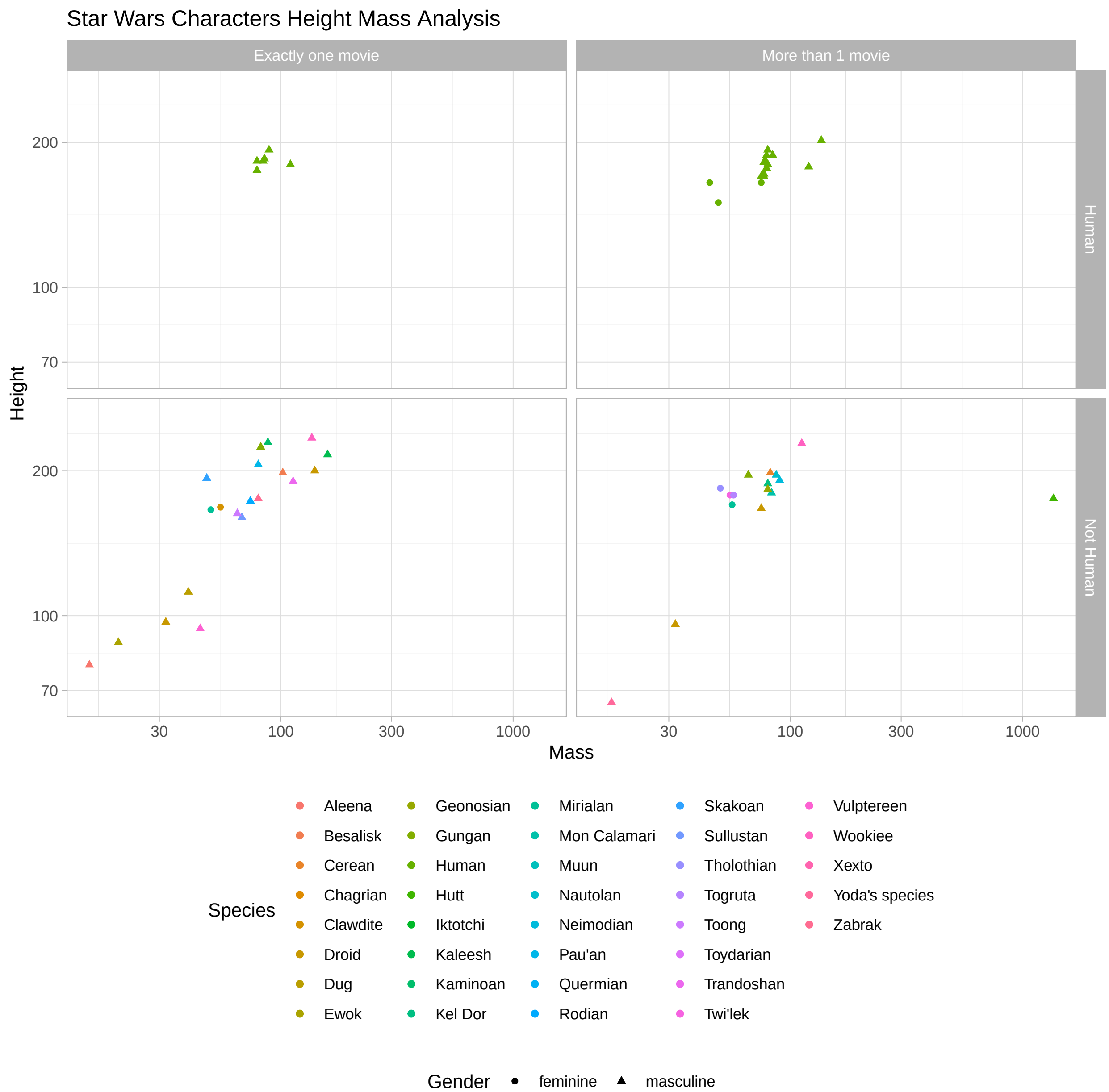
1. What is Body - Mass distribution of characters per species and genders?
   I want to see if there are some stereotypes for each species and genders.
2. What is female characters' rate among all?
   I think number of feminine characters is much lower than the one of masculine characters.
3. Whether, out of range characters are more common in whole movies or special to one movie only?
   It can show us if more uncommon characters are seen only in one movie or not.

Outcomes for there 3 questions are as below :

1. There is a clear stereotype for human characters. Height is around 1.70 - 1.90 meters and mass is around 75 - 100 kilograms. Feminine characters are outliers of that stereotype. Only one feminine observation is above 1.70 m 70 kg mass.
   Height and mass are more dispersed for non-human species. Still, many characters between 1.70 m and 2 m tall and 65 - 100 kg weight. Outliers of this group are mostly observed in masculine characters.
2. Only 17 of 87 characters are feminine. 8 of them are showing up only 1 movie. At the end we have only 7 permanent feminine characters. In a movie series having different worlds, species and time periods, that rate is so ridiculous. That may be caused of the genre of the movies. Producers might be thinking the audience of sci-fi movies will mostly be males, but it is ended up differently. That idea can also be supported by the comments about women characters being over sexualized in the series.
3. Smaller characters in terms of both height and weight are seen among non-human characters and they are showing up mostly for one movie only. Same is seen for taller characters as well, they are mostly non-human and seen for one movie only. It can be said that creators bought some odd characters special for some movies but permanent characters have some sense of body range.

```
starwars %>%
  rowwise() %>%
  mutate(n_films = length(films)) %>%
  mutate(more_1 = case_when(n_films == 1 ~ "Exactly one movie",
```

```
                              n_films != 1 ~ "More than 1 movie")) %>%
  mutate(human = case_when(species == "Human" ~ "Human",
                             species != "Human" ~ "Not Human")) %>%
  filter(gender %in% c("feminine", "masculine"), !is.na(human)) %>%
  ggplot(aes(mass,height)) +
  facet_grid(human ~ more_1) +
  geom_point(aes(col=species,shape=gender)) +
  scale_y_log10() +
  scale_x_log10() +
  ggtitle("Star Wars Characters Height Mass Analysis") +
  labs(x = "Mass",
       y = "Height",
       col="Species",
       shape="Gender") +

  theme_light() +
  theme(legend.position = "bottom",
        legend.box = "vertical")
```



Part II: Extending Group Project

Our project was about marriage statistics. We worked on number of first marriages through years. Analysis is made per gender, age group and education level.

The group project has 2 shortfalls :

- It isn't include ratio over population. (For Age Groups and Education Levels)
- We couldn't show first marriage rates of group among others. (For Age Groups and Education Levels)

Thus, I want to show ratio information on top of our project.

1. First Marriage Rate Trend Over Population Data

1.a Data Preparation I found population data with age information from TUIK database and prepared the data for analysis purpose. That data doesn't have an age group as "16-19" so, I filtered age groups from 20 up to 59 years old (included). 60+ population is ignored as first marriage data has a very low portion in that group.
Longer pivot format is implemented for the analysis and number of first marriages of same age groups are merged with the population information. First marriage rate is calculated by dividing the first marriage number to the population size.

```
##Import Project Data From Our GitHub Page
Final=tempfile(fileext=".RDS")
download.file("https://github.com/pjournal/boun01g-r-ammstein/blob/gh-pages/FinalProjectRawData.RDS?raw=
              destfile=Final,mode='wb')
FinalRawData = readRDS(Final)
file.remove(Final)
```

```
## [1] TRUE
```

```
##Import Population Data From TUIK
PopulationRawData  <-
  read.csv("C:/Users/ecemnaz.bay/Desktop/Ecemnaz/Master/IE 48A/boun01-Ecemnaz0/NufusVerisi.csv")

##Rename Columns
PopulationRawData <- PopulationRawData %>%
  rename( "Years"="ï..Yillar"
 ,"Male - 0-4"="Erkek.ve.0.4"
 ,"Male - 10-14"="Erkek.ve.10.14"
 ,"Male - 15-19"="Erkek.ve.15.19"
 ,"Male - 20-24"="Erkek.ve.20.24"
 ,"Male - 25-29"="Erkek.ve.25.29"
 ,"Male - 30-34"="Erkek.ve.30.34"
 ,"Male - 35-39"="Erkek.ve.35.39"
 ,"Male - 40-44"="Erkek.ve.40.44"
 ,"Male - 45-49"="Erkek.ve.45.49"
 ,"Male - 50-54"="Erkek.ve.50.54"
 ,"Male - 55-59"="Erkek.ve.55.59"
 ,"Male - 5-9"="Erkek.ve.5.9"
 ,"Male - 60-64"="Erkek.ve.60.64"
 ,"Male - 65-69"="Erkek.ve.65.69"
 ,"Male - 70-74"="Erkek.ve.70.74"
 ,"Male - 75-79"="Erkek.ve.75.79"
 ,"Male - 80-84"="Erkek.ve.80.84"
 ,"Male - 85-89"="Erkek.ve.85.89"
 ,"Male - 90+"="Erkek.ve.90."
 ,"Female - 0-4"="KadÄ.n.ve.0.4"
 ,"Female - 10-14"="KadÄ.n.ve.10.14"
 ,"Female - 15-19"="KadÄ.n.ve.15.19"
 ,"Female - 20-24"="KadÄ.n.ve.20.24"
 ,"Female - 25-29"="KadÄ.n.ve.25.29"
 ,"Female - 30-34"="KadÄ.n.ve.30.34"
 ,"Female - 35-39"="KadÄ.n.ve.35.39"
 ,"Female - 40-44"="KadÄ.n.ve.40.44"
 ,"Female - 45-49"="KadÄ.n.ve.45.49"
 ,"Female - 50-54"="KadÄ.n.ve.50.54"
 ,"Female - 55-59"="KadÄ.n.ve.55.59"
```

```
,"Female - 5-9"="KadÄ.n.ve.5.9"
,"Female - 60-64"="KadÄ.n.ve.60.64"
,"Female - 65-69"="KadÄ.n.ve.65.69"
,"Female - 70-74"="KadÄ.n.ve.70.74"
,"Female - 75-79"="KadÄ.n.ve.75.79"
,"Female - 80-84"="KadÄ.n.ve.80.84"
,"Female - 85-89"="KadÄ.n.ve.85.89"
,"Female - 90+"="KadÄ.n.ve.90.")

PopulationRawData <- drop_na(PopulationRawData)

PopulationRawData_Longer <-
PopulationRawData%>% pivot_longer(cols=c(-Years))

PopulationRawData_Longer_NewColumns <-
separate( PopulationRawData_Longer
         ,name
         ,into = c("Gender","AgeGroup")
         ,sep = "-"
         ,remove = TRUE
         ,extra = "merge")

PopulationRawData_Longer_Fitered <-
PopulationRawData_Longer_NewColumns %>%
 mutate(AgeGroup=trimws(PopulationRawData_Longer_NewColumns$AgeGroup,which="both")) %>%
 mutate(Gender=trimws(PopulationRawData_Longer_NewColumns$Gender,which="both")) %>%
 rename("Year"="Years") %>%
 rename("Population"="value") %>%
 filter(AgeGroup %in% c("20-24","25-29","30-34","35-39","40-44","45-49","50-54","55-59") )

PopulationRawData_Longer_Fitered$AgeGroup <-
  as.factor(PopulationRawData_Longer_Fitered$AgeGroup)
PopulationRawData_Longer_Fitered$Gender <-
  as.factor(PopulationRawData_Longer_Fitered$Gender)

FinalData_AgeBased <- FinalRawData %>%
  filter(AgeGroup %in% c("20-24","25-29","30-34","35-39"
                        ,"40-44","45-49","50-54","55-59") ) %>%
  group_by(Year,AgeGroup,Gender) %>%
  summarise(NbOfFirstMarriages=sum(NbOfFirstMarriages))

FinalData_AgeBased$Year <-
  as.integer(FinalData_AgeBased$Year)
FinalData_AgeBased$NbOfFirstMarriages <-
  as.integer(FinalData_AgeBased$NbOfFirstMarriages)

Marriage_Population_Joined <-
  left_join(FinalData_AgeBased, PopulationRawData_Longer_Fitered)

Marriage_Population_Rate <- Marriage_Population_Joined %>%
  mutate(NbOfFirstMarriages/Population)

Female_Marriage_Population_Rate <- Marriage_Population_Rate %>%
```

```
  filter(Gender == "Female")

Male_Marriage_Population_Rate <- Marriage_Population_Rate %>%
  filter(Gender == "Male")
```

1.b First Marriage Rate Trend
Female Rates

- Rate of women having their first marriage at ages of 20-24 is decreasing since 2015.
- Rate of women having their first marriage at ages of 25-29 is increasing since 2015.
- Rate of higher age gorups doesn't seem to be changing much.

Male Rates

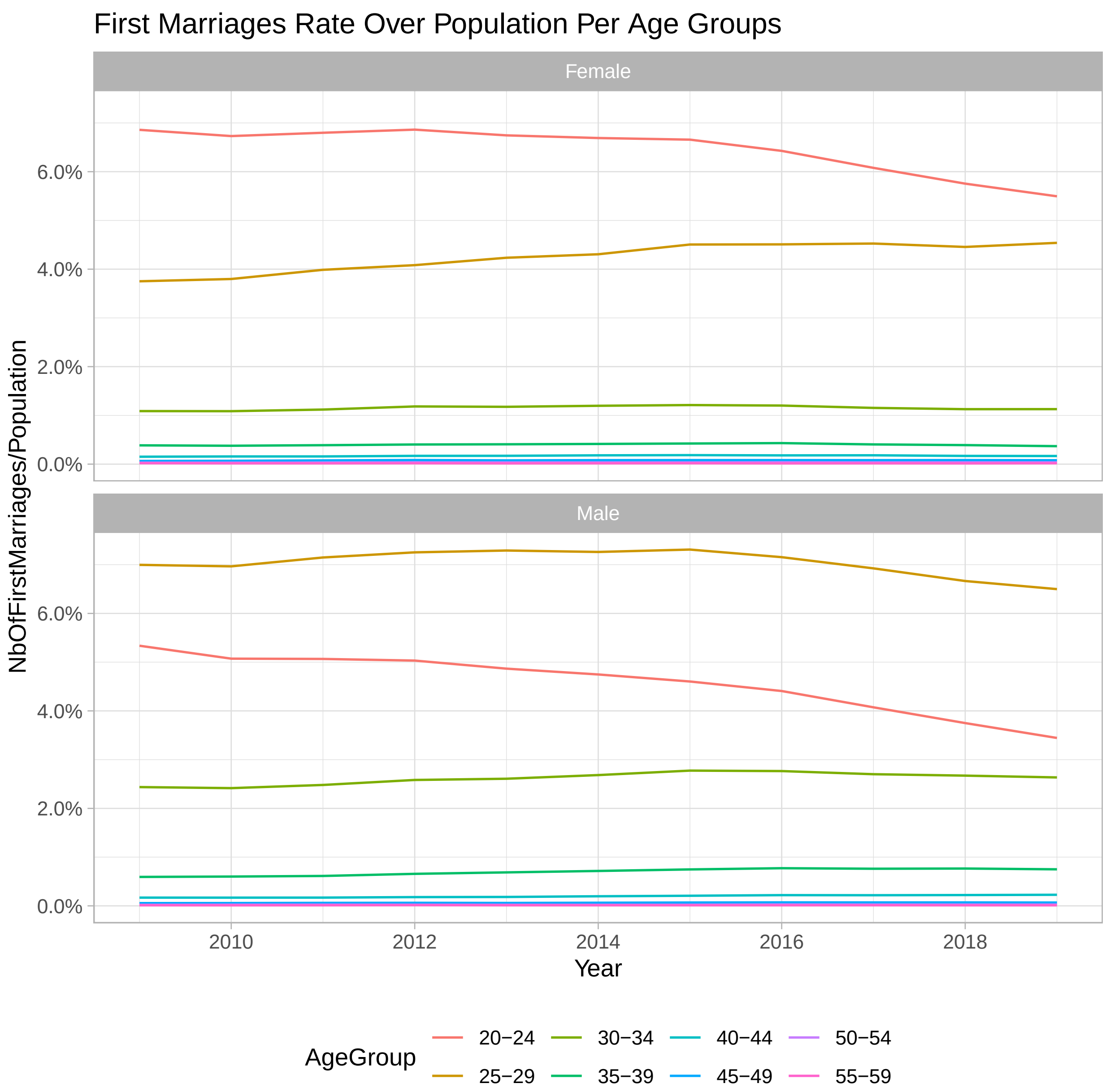
- Rates of men having their first marriage at ages of 20-24 and 25-29 are decreasing since 2015.
- Rate of men having their first marriage at ages of 30-34 is increasing since 2015.
- Rate of higher age gorups doesn't seem to be changing much.

Comparisons

- First marriage rate for 30-34 age group is significantly higher for men than women.
- The highest rates are seen at 25-29 age group for men, and 20-24 age group for women despite the changes.
- Decrease of rate for 20-24 age group is more apparent for men.
- Only clear increase is seen for 25-29 age group's rate for women.

Those findings also supports the outcomes we got from the analysis in the project. Education level information can be implemented into that analysis as well.

```
ggplot(Marriage_Population_Rate,aes(x=Year,y=NbOfFirstMarriages/Population,color=AgeGroup)) +
  geom_line() +
  facet_wrap(~Gender,dir = "v") +
  theme_light() +
  theme(legend.position = "bottom",
        legend.box = "vertical") +
  scale_x_continuous(breaks = pretty_breaks()) +
  scale_y_continuous(labels = scales::percent) +
  ggtitle("First Marriages Rate Over Population Per Age Groups")
```



2. First Marriage Rates Comparison Among Age Groups

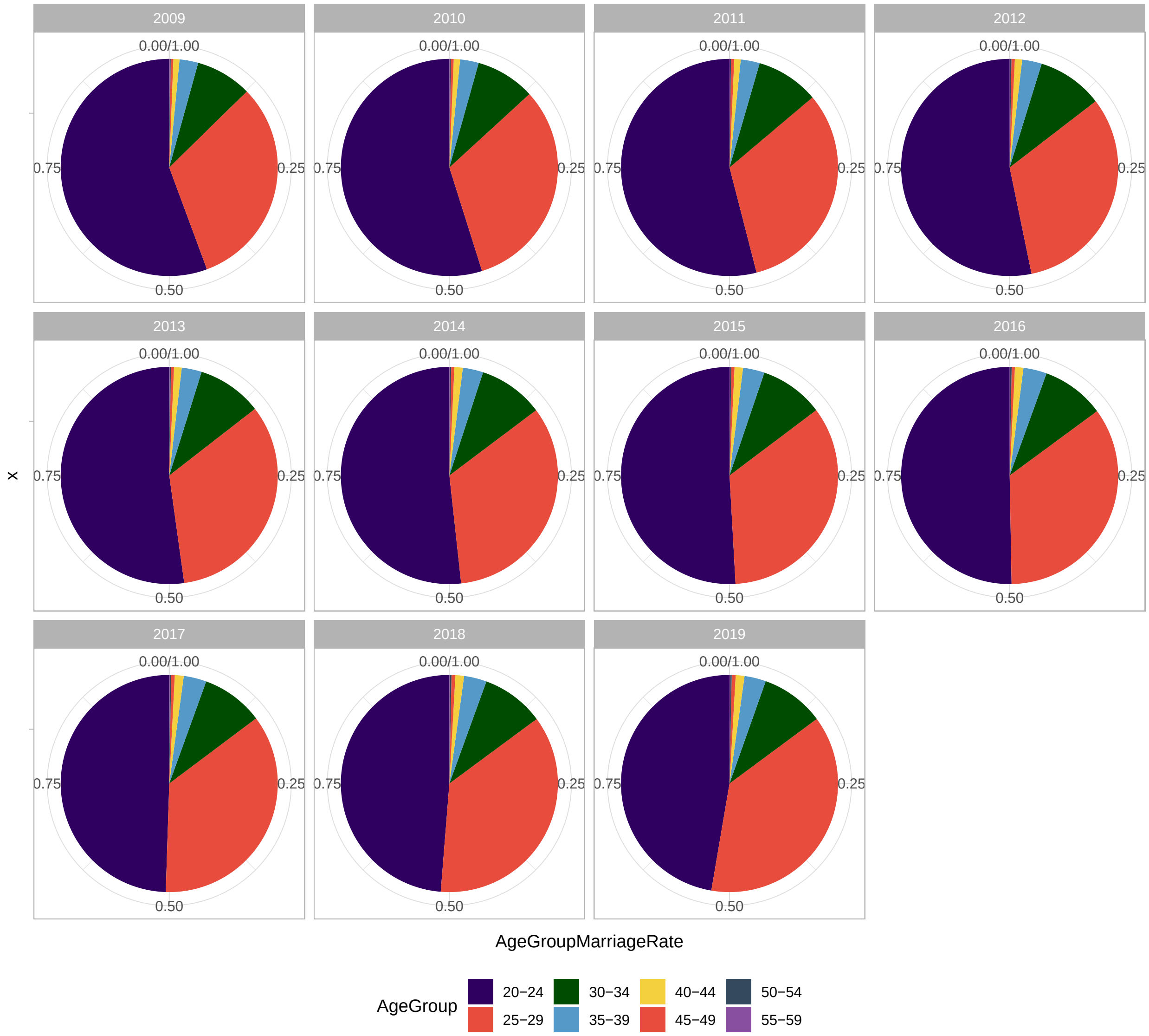
Below pie charts show the distribution of first marriage numbers among age groups per year.
Correlation with he rate trend in previous part is significant. In the current years,

- The share of 20-24 age group for women is shrinking, while the one for 25-29 age group is expanding.
- The share of 30-34 age group for men is expanding.

```
##Female First Marriage Rate By Age Group
Marriage_Population_Rate_Female <- Marriage_Population_Rate %>%
  filter(Gender=="Female")

Marriage_Population_Rate_Female_Totals <- Marriage_Population_Rate_Female %>%
  group_by(Year,Gender) %>%
```

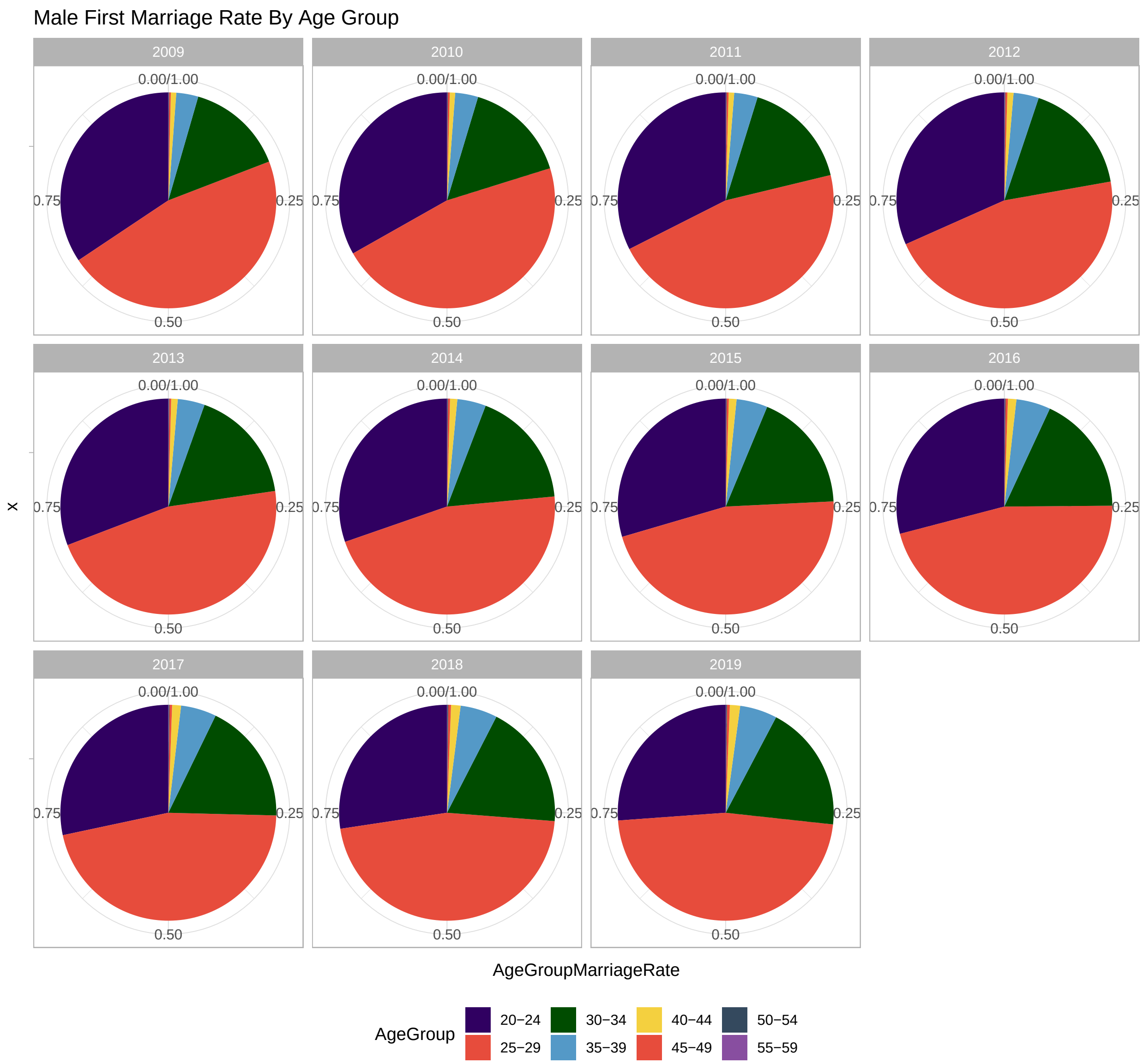
```
  summarise(NbOfFirstMarriages_YearTotal=sum(NbOfFirstMarriages))

Female_Age_MarriageRate <-
  left_join(Marriage_Population_Rate_Female,Marriage_Population_Rate_Female_Totals) %>%
  select(-c("NbOfFirstMarriages/Population","Population")) %>%
  mutate(AgeGroupMarriageRate=(NbOfFirstMarriages/NbOfFirstMarriages_YearTotal))


ggplot(Female_Age_MarriageRate) +
  facet_wrap(~Year) +
  geom_bar(aes(y = AgeGroupMarriageRate, x = "", fill = AgeGroup), stat = "identity") +
  theme() +
  coord_polar("y", start = 0)+
  theme_light() +
  theme(legend.position = "bottom",
        legend.box = "vertical") +
  scale_fill_manual(values=c("#300061","#E74C3C","#004C00",
                             "#5499C7", "#F4D03F", "#E74C3C",
                             "#34495E","#884EA0")) +
  ggtitle("Female First Marriage Rate By Age Group")
```

Female First Marriage Rate By Age Group




```
##Male First Marriage Rate By Age Group

Marriage_Population_Rate_Male <- Marriage_Population_Rate %>%
  filter(Gender=="Male")

Marriage_Population_Rate_Male_Totals <- Marriage_Population_Rate_Male %>%
  group_by(Year,Gender) %>%
  summarise(NbOfFirstMarriages_YearTotal=sum(NbOfFirstMarriages))

Male_Age_MarriageRate <-
  left_join(Marriage_Population_Rate_Male,Marriage_Population_Rate_Male_Totals) %>%
  select(-c("NbOfFirstMarriages/Population","Population")) %>%
  mutate(AgeGroupMarriageRate=(NbOfFirstMarriages/NbOfFirstMarriages_YearTotal))
```

```
ggplot(Male_Age_MarriageRate) +
  facet_wrap(~Year) +
  geom_bar(aes(y = AgeGroupMarriageRate, x = "", fill = AgeGroup), stat = "identity") +
  theme() +
  coord_polar("y", start = 0)+
  theme_light() +
  theme(legend.position = "bottom",
          legend.box = "vertical") +
  scale_fill_manual(values=c("#300061","#E74C3C","#004C00",
                             "#5499C7", "#F4D03F", "#E74C3C",
                             "#34495E","#884EA0")) +
  ggtitle("Male First Marriage Rate By Age Group")
```



Part III: Real Life Example

1. Data Preparation

2 years data combined from ODD website (2018 August-2020 August). We worked with my group member Canan for gathering data. Null values are replaced with 0, as no sales are seen in that period for that brand.

```
ODDRawData <-
  readRDS("C:/Users/ecemnaz.bay/Desktop/Ecemnaz/Master/IE 48A/boun01-Ecemnaz0/OddRawData.RDS")

ODDRawData <- ODDRawData %>%
  replace(is.na(.), 0) %>%
  select(-year_month) %>%
  mutate(YearMonth=(100*year)+month)

ODDRawData_Longer <- ODDRawData %>%
  pivot_longer(cols=c(oto_dom,oto_imp,lcv_dom,lcv_imp))

DomesticAutom_MonthlyTotalSales <- ODDRawData_Longer %>%
  filter(name=="oto_dom") %>%
  group_by(YearMonth) %>%
  summarise(MonthlyTotalSum=sum(value))

DomesticAutom_RawData <- ODDRawData_Longer %>%
  filter(name=="oto_dom")

ImportedAutom_MonthlyTotalSales <- ODDRawData_Longer %>%
  filter(name=="oto_imp") %>%
  group_by(YearMonth) %>%
  summarise(MonthlyTotalSum=sum(value))

ImportedAutom_RawData <- ODDRawData_Longer %>%
  filter(name=="oto_imp")

DomesticLighhtCV_MonthlyTotalSales <- ODDRawData_Longer %>%
  filter(name=="lcv_dom") %>%
  group_by(YearMonth) %>%
  summarise(MonthlyTotalSum=sum(value))

DomesticLighhtCV_RawData <- ODDRawData_Longer %>%
  filter(name=="lcv_dom")

ImportedLighhtCV_MonthlyTotalSales <- ODDRawData_Longer %>%
  filter(name=="lcv_imp") %>%
  group_by(YearMonth) %>%
  summarise(MonthlyTotalSum=sum(value))

ImportedLighhtCV_RawData <- ODDRawData_Longer %>%
  filter(name=="lcv_imp")

DomesticAutom_MonthlyTotalSalesRates <-
  left_join(DomesticAutom_RawData,DomesticAutom_MonthlyTotalSales) %>%
  mutate(MonthlySalesRate=percent(value/MonthlyTotalSum,0.01)) %>%
  arrange(desc(value))
```

```r
ImportedAutom_MonthlyTotalSalesRates <-
  left_join(ImportedAutom_RawData,ImportedAutom_MonthlyTotalSales) %>%
  mutate(MonthlySalesRate=percent(value/MonthlyTotalSum,0.01)) %>%
  arrange(desc(value))

DomesticLighhtCV_MonthlyTotalSalesRates <-
  left_join(DomesticLighhtCV_RawData,DomesticLighhtCV_MonthlyTotalSales) %>%
  mutate(MonthlySalesRate=percent(value/MonthlyTotalSum,0.01)) %>%
  arrange(desc(value))

ImportedLighhtCV_MonthlyTotalSalesRates <-
  left_join(ImportedLighhtCV_RawData,ImportedLighhtCV_MonthlyTotalSales) %>%
  mutate(MonthlySalesRate=percent(value/MonthlyTotalSum,0.01)) %>%
  arrange(desc(value))

DomesticAutomSalesTrendData <- DomesticAutom_MonthlyTotalSalesRates %>%
                                filter(year>2018)%>%
                                group_by(YearMonth) %>%
                                top_n(10,value)

ImportedAutomSalesTrendData <- ImportedAutom_MonthlyTotalSalesRates %>%
                                filter(year>2018)%>%
                                group_by(YearMonth) %>%
                                top_n(10,value)

DomesticLightCVSalesTrendData <- DomesticLighhtCV_MonthlyTotalSalesRates %>%
                                  filter(year>2018)%>%
                                  group_by(YearMonth) %>%
                                  top_n(10,value)

ImportedLightCVSalesTrendData <- ImportedLighhtCV_MonthlyTotalSalesRates %>%
                                  filter(year>2018)%>%
                                  group_by(YearMonth) %>%
                                  top_n(10,value)


DomesticAutomSalesTrend <- ggplot( DomesticAutomSalesTrendData,
        aes(x=month,y=value,color=brand)) +
        facet_wrap(~year) +
        geom_col(aes(fill=brand)) +
        scale_x_continuous(breaks = pretty_breaks(),limits = c(0,9)) +
        ggtitle("Domestic Automobile") +
        labs(x = "Month",
             y = "Sales Quantity") +
        theme_light() +
        theme(legend.position = "bottom",legend.box = "vertical",
              legend.title = element_blank())

ImportedAutomSalesTrend <- ggplot( ImportedAutomSalesTrendData,
        aes(x=month,y=value,color=brand)) +
        facet_wrap(~year) +
        geom_col(aes(fill=brand)) +
        scale_x_continuous(breaks = pretty_breaks(),limits = c(0,9)) +
```

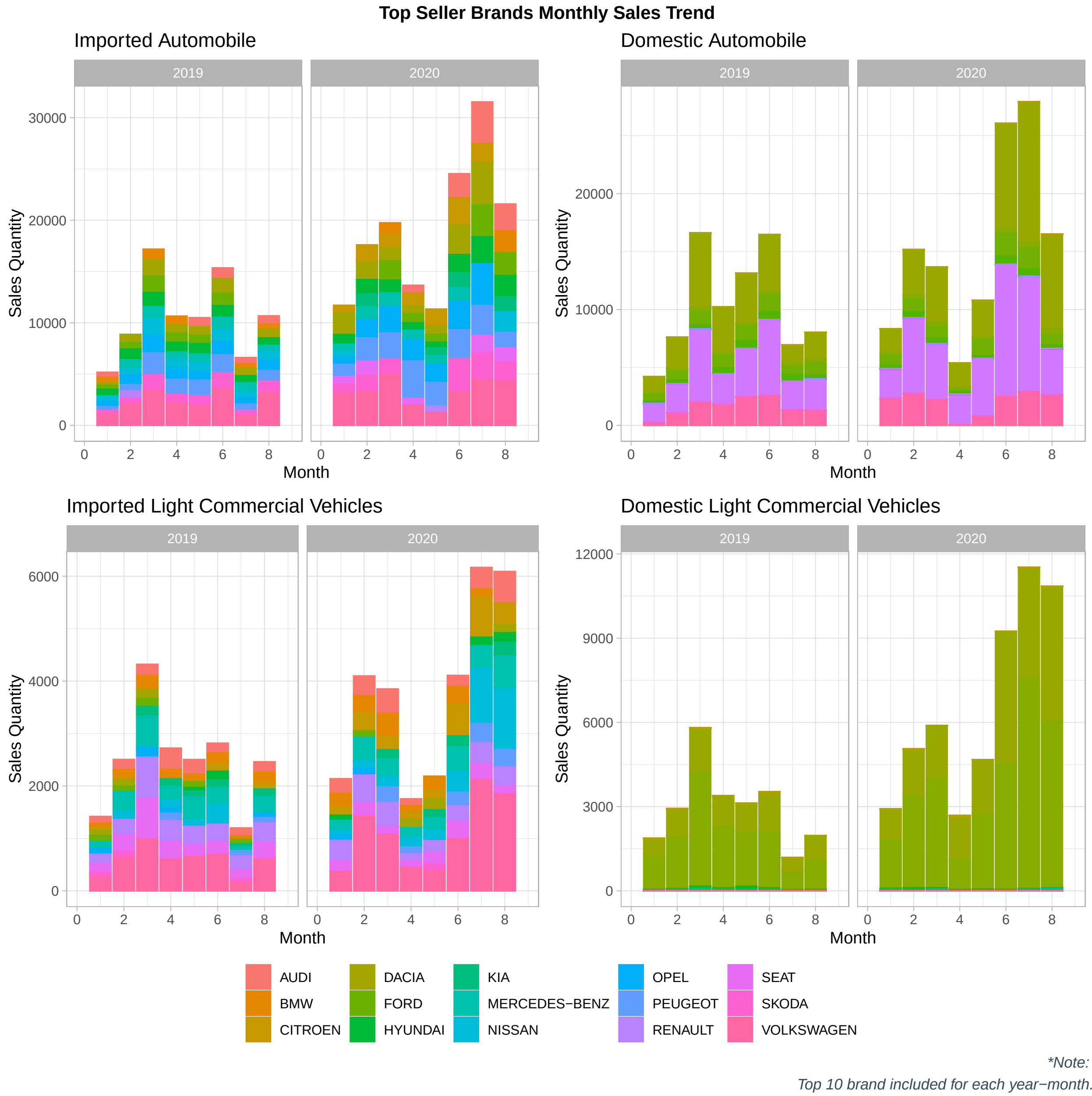
```r
        ggtitle("Imported Automobile") +
        labs(x = "Month",
             y = "Sales Quantity") +
        theme_light() +
        theme(legend.position = "bottom",legend.box = "vertical",
              legend.title = element_blank())

DomesticLightCVSalesTrend <- ggplot( DomesticLightCVSalesTrendData,
        aes(x=month,y=value,color=brand)) +
        facet_wrap(~year) +
        geom_col(aes(fill=brand)) +
        scale_x_continuous(breaks = pretty_breaks(),limits = c(0,9)) +
        ggtitle("Domestic Light Commercial Vehicles") +
        labs(x = "Month",
             y = "Sales Quantity") +
        theme_light() +
        theme(legend.position = "bottom",legend.box = "vertical",
              legend.title = element_blank())

ImportedLightCVSalesTrend <- ggplot( ImportedLightCVSalesTrendData,
        aes(x=month,y=value,color=brand)) +
        facet_wrap(~year) +
        geom_col(aes(fill=brand)) +
        scale_x_continuous(breaks = pretty_breaks(),limits = c(0,9)) +
        ggtitle("Imported Light Commercial Vehicles") +
        labs(x = "Month",
             y = "Sales Quantity") +
        theme_light() +
        theme(legend.position = "bottom",legend.box = "vertical",
              legend.title = element_blank())

figure <-  ggarrange(ImportedAutomSalesTrend,
            DomesticAutomSalesTrend,
            ImportedLightCVSalesTrend,
            DomesticLightCVSalesTrend,
            common.legend = TRUE, legend = "bottom",ncol = 2, nrow = 2)
```

2. Top Sellers Sales Trend

Top 10 brands are selected for each year-week. Year to date comparisons of 2019 vs 2020 is observed as below figure.

- Overall demand is increased for automobiles. That trend is expected as private cars are much more preferable than public transportation this year, due to pandemic.
- Brand dominance doesn't seem to be changing much, except the increase of Audi sales.
- Same comments are also applicable for light commercial vehicles despite the fact that there are only a small number of domestic brands.

```r
annotate_figure(figure,
               top = text_grob("Top Seller Brands Monthly Sales Trend",
                               face = "bold", size = 12),
               bottom = text_grob("*Note: \n Top 10 brand included for each year-month.",
                                  color = "#34495E",
                                  hjust = 1, x = 1, face = "italic", size = 10))
```



3. Best Sellers Through Years

Table 1 shows that since the beginning of 2019, Renault and Fiat have the largest sales rate among other brands. People tend to buy cheaper segments if they buy domestic cars.
Table 2 shows that Volkswagen is dominating the imported car market. That brand is highly preferred by car fleets. That may be effecting the sales rate.

```
DomesticAutom_MonthlyTotalSalesRates %>%
  rename(Brand=brand) %>%
  group_by(YearMonth) %>%
  top_n(1,value) %>%
  arrange(desc(YearMonth)) %>%
  select(Brand,YearMonth,MonthlySalesRate) %>%
  kbl(caption = "Domestic Automobile Sales Rate") %>%
```

Table 1: Domestic Automobile Sales Rate

Brand	YearMonth	MonthlySalesRate
FIAT	202008	49.48%
FIAT	202007	43.57%
RENAULT	202006	43.70%
RENAULT	202005	46.23%
RENAULT	202004	48.79%
RENAULT	202003	35.15%
RENAULT	202002	43.04%
RENAULT	202001	30.72%
FIAT	201912	42.79%
FIAT	201911	41.46%
FIAT	201910	33.28%
RENAULT	201909	39.68%
RENAULT	201908	33.78%
RENAULT	201907	35.73%
RENAULT	201906	39.88%
FIAT	201905	33.06%
FIAT	201904	38.26%
FIAT	201903	38.88%
FIAT	201902	35.65%
RENAULT	201901	39.66%
FIAT	201812	36.92%
TOYOTA	201811	33.13%
HONDA	201810	29.72%
RENAULT	201809	36.10%
RENAULT	201808	30.90%

```
  kable_minimal(full_width = F)
```

```
ImportedAutom_MonthlyTotalSalesRates %>%
  rename(Brand=brand) %>%
  group_by(YearMonth) %>%
  top_n(1,value) %>%
  arrange(desc(YearMonth)) %>%
  select(Brand,YearMonth,MonthlySalesRate) %>%
  kbl(caption = "Imported Automobile Sales Rate") %>%
  kable_minimal(full_width = F)
```

Table 2: Imported Automobile Sales Rate

Brand	YearMonth	MonthlySalesRate
VOLKSWAGEN	202008	15.99%
VOLKSWAGEN	202007	11.05%
VOLKSWAGEN	202006	11.10%
PEUGEOT	202005	17.30%
PEUGEOT	202004	23.06%
VOLKSWAGEN	202003	19.68%
VOLKSWAGEN	202002	15.37%
VOLKSWAGEN	202001	24.05%
VOLKSWAGEN	201912	14.38%
PEUGEOT	201911	13.61%
VOLKSWAGEN	201910	19.01%
VOLKSWAGEN	201909	19.38%
VOLKSWAGEN	201908	25.01%
VOLKSWAGEN	201907	13.06%
VOLKSWAGEN	201906	18.74%
VOLKSWAGEN	201905	14.96%
VOLKSWAGEN	201904	16.42%
VOLKSWAGEN	201903	16.01%
VOLKSWAGEN	201902	19.10%
VOLKSWAGEN	201901	17.86%
VOLKSWAGEN	201812	17.50%
NISSAN	201811	14.33%
NISSAN	201810	14.31%
VOLKSWAGEN	201809	14.52%
VOLKSWAGEN	201808	13.81%