R Project 2

EDA With R-Banking Data Set

**Task**

1. Explore the data to find how different features (age, job, education, and others) affect the desired outcome (the client subscribed to a term deposit). For this analysis, I will use a marketing KPI called \*Conversion Rate\*. Conversion rate is the percentage of clients who take the desired action.

2. Give recommendations for the Bank's marketing strategy and future marketing campaigns.

**## \*\*Loading the data and R packages\*\***

```{r}

library(dplyr)

library(ggplot2)

data <- read.csv("C:/Users/LENOVO/Downloads/bank-additional-full(1).csv",

sep = ";", header = TRUE)

head(data)

```

| **age**  <int> | **job**  <chr> | **marital**  <chr> | **education**  <chr> | **default**  <chr> | **housing**  <chr> | **loan**  <chr> | **contact**  <chr> | **month**  <chr> | **day\_of\_week**  <chr> | **duration**  <int> | **campaign**  <int> | **pdays**  <int> | **previous**  <int> | **poutcome**  <chr> |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | 56 | Housemaid | married | basic.4y | no | no | no | telephone | may | mon | 261 | 1 | 999 | 0 | nonexistent |  |
| 2 | 57 | Services | married | high.school | unknown | no | no | telephone | may | mon | 149 | 1 | 999 | 0 | nonexistent |  |
| 3 | 37 | Services | married | high.school | no | yes | no | telephone | may | mon | 226 | 1 | 999 | 0 | nonexistent |  |
| 4 | 40 | admin. | married | basic.6y | no | no | no | telephone | may | mon | 151 | 1 | 999 | 0 | nonexistent |  |
| 5 | 56 | Services | married | high.school | no | no | yes | telephone | may | mon | 307 | 1 | 999 | 0 | nonexistent |  |
| 6 | 45 | Services | married | basic.9y | unknown | no | no | telephone | may | mon | 198 | 1 | 999 | 0 | nonexistent |  |

6 rows | 1-16 of 21 columns

| **month**  <chr> |  | **day\_of\_week**  <chr> | **duration**  <int> | **campaign**  <int> | **pdays**  <int> | **previous**  <int> | **poutcome**  <chr> | **emp.var.rate**  <dbl> | **cons.price.idx**  <dbl> | **cons.conf.idx**  <dbl> | **euribor3m**  <dbl> | **nr.employed**  <dbl> | **y**  <chr> |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | may | mon | 261 | 1 | 999 | 0 | nonexistent | 1.1 | 93.994 | -36.4 | 4.857 | 5191 | no |
|  |  | may | mon | 149 | 1 | 999 | 0 | nonexistent | 1.1 | 93.994 | -36.4 | 4.857 | 5191 | no |
|  |  | may | mon | 226 | 1 | 999 | 0 | nonexistent | 1.1 | 93.994 | -36.4 | 4.857 | 5191 | no |
|  |  | may | mon | 151 | 1 | 999 | 0 | nonexistent | 1.1 | 93.994 | -36.4 | 4.857 | 5191 | no |
|  |  | may | mon | 307 | 1 | 999 | 0 | nonexistent | 1.1 | 93.994 | -36.4 | 4.857 | 5191 | no |
|  |  | may | mon | 198 | 1 | 999 | 0 | nonexistent | 1.1 | 93.994 | -36.4 | 4.857 | 5191 | no |

6 rows | 10-22 of 21 columns

> library(dplyr)

> library(ggplot2)

> data <- read.csv("C:/Users/LENOVO/Downloads/bank-additional-full(1).csv",

+ sep = ";", header = TRUE)

>

> head(data)

> data <- data %>%

+ mutate(y=ifelse(y=="no", 0, 1))

> data$y <- as.integer(data$y)

```{r}

data <- data %>%

mutate(y=ifelse(y=="no", 0, 1))

data$y <- as.integer(data$y)

#total number of conversions

sum(data$y)

#total number of clients in the data

nrow(data)

#conversion rate

sum(data$y)/nrow(data)\*100.0

```

[1] 4640

[1] 41188

[1] 11.26542

> #total number of conversions

> sum(data$y)

[1] 4640

>

> #total number of clients in the data

> nrow(data)

[1] 41188

>

> #conversion rate

> sum(data$y)/nrow(data)\*100.0

[1] 11.26542

> data <- data %>%

+ mutate(y=ifelse(y=="no", 0, 1))

> data$y <- as.integer(data$y)

>

> #total number of conversions

> sum(data$y)

[1] 41188

>

> #total number of clients in the data

> nrow(data)

[1] 41188

>

> #conversion rate

> sum(data$y)/nrow(data)\*100.0

[1] 100

## \*\*Conversion Rate by Age\*\*

```{r}

#group clients into 6 age groups(18-30, 30-40, 40-50, 50-60, 60-70, >70)

conversionsAgeGroup <- data %>%

group\_by(AgeGroup=cut(age, breaks=seq(20, 70, by=10))) %>%

summarize(TotalCount=n(), NumberConversions=sum(y)) %>%

mutate(ConversionRate=NumberConversions/TotalCount\*100)

#rename the 6th group

conversionsAgeGroup$AgeGroup <- as.character(conversionsAgeGroup$AgeGroup)

conversionsAgeGroup$AgeGroup[6] <- "70+"

#visualizing conversions by age group

ggplot(data=conversionsAgeGroup, aes(x=AgeGroup, y=ConversionRate)) +

geom\_bar(width=0.5, stat="identity", fill="darkgreen") +

labs(title="Conversion Rates by Age Group")

```

> #group clients into 6 age groups(18-30, 30-40, 40-50, 50-60, 60-70, >70)

> conversionsAgeGroup <- data %>%

+ group\_by(AgeGroup=cut(age, breaks=seq(20, 70, by=10))) %>%

+ summarize(TotalCount=n(), NumberConversions=sum(y)) %>%

+ mutate(ConversionRate=NumberConversions/TotalCount\*100)

>

> #rename the 6th group

> conversionsAgeGroup$AgeGroup <- as.character(conversionsAgeGroup$AgeGroup)

> conversionsAgeGroup$AgeGroup[6] <- "70+"

>

> #visualizing conversions by age group

> ggplot(data=conversionsAgeGroup, aes(x=AgeGroup, y=ConversionRate)) +

+ geom\_bar(width=0.5, stat="identity", fill="darkgreen") +

+ labs(title="Conversion Rates by Age Group")

> #group clients into 6 age groups(18-30, 30-40, 40-50, 50-60, 60-70, >70)

> conversionsAgeGroup <- data %>%

+ group\_by(AgeGroup=cut(age, breaks=seq(20, 70, by=10))) %>%

+ summarize(TotalCount=n(), NumberConversions=sum(y)) %>%

+ mutate(ConversionRate=NumberConversions/TotalCount\*100)

>

> #rename the 6th group

> conversionsAgeGroup$AgeGroup <- as.character(conversionsAgeGroup$AgeGroup)

> conversionsAgeGroup$AgeGroup[6] <- "70+"

>

> #visualizing conversions by age group

> ggplot(data=conversionsAgeGroup, aes(x=AgeGroup, y=ConversionRate)) +

+ geom\_bar(width=0.5, stat="identity", fill="darkgreen") +

+ labs(title="Conversion Rates by Age Group")

> # group the data

> conversionsAgeMarital <- data %>%

+ group\_by(AgeGroup=cut(age, breaks=seq(20,70, by=10)),

+ Marital=marital) %>%

+ summarize(Count=n(), NumConversions=sum(y)) %>%

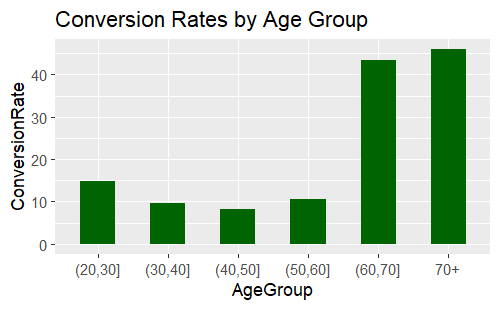
+ mutate(TotalCount=sum(Count)) %>%

+ mutate(ConversionRate=NumConversions/TotalCount\*100)

`summarise()` has grouped output

by 'AgeGroup'. You can override

using the `.groups` argument.



## \*\*Conversions by age group and marital status\*\*

```{r}

# group the data

conversionsAgeMarital <- data %>%

group\_by(AgeGroup=cut(age, breaks=seq(20,70, by=10)),

Marital=marital) %>%

summarize(Count=n(), NumConversions=sum(y)) %>%

mutate(TotalCount=sum(Count)) %>%

mutate(ConversionRate=NumConversions/TotalCount\*100)

#rename the last groups

conversionsAgeMarital$AgeGroup <- as.character(conversionsAgeMarital$AgeGroup)

conversionsAgeMarital$AgeGroup[is.na(conversionsAgeMarital$AgeGroup)] <- "70+"

#visualizing conversions by age group and marrital status

ggplot(conversionsAgeMarital, aes(x=AgeGroup, y=ConversionRate, fill=Marital)) +

geom\_bar(width=0.5, stat = "identity") +

labs(title="Conversion Rates by Age Group and Marital Status")

```

> #rename the last groups

> conversionsAgeMarital$AgeGroup <- as.character(conversionsAgeMarital$AgeGroup)

> conversionsAgeMarital$AgeGroup[is.na(conversionsAgeMarital$AgeGroup)] <- "70+"

>

> #visualizing conversions by age group and marrital status

> ggplot(conversionsAgeMarital, aes(x=AgeGroup, y=ConversionRate, fill=Marital)) +

+ geom\_bar(width=0.5, stat = "identity") +

+ labs(title="Conversion Rates by Age Group and Marital Status")

> # group the data

> conversionsAgeMarital <- data %>%

+ group\_by(AgeGroup=cut(age, breaks=seq(20,70, by=10)),

+ Marital=marital) %>%

+ summarize(Count=n(), NumConversions=sum(y)) %>%

+ mutate(TotalCount=sum(Count)) %>%

+ mutate(ConversionRate=NumConversions/TotalCount\*100)

`summarise()` has grouped output

by 'AgeGroup'. You can override

using the `.groups` argument.

.

> #rename the last groups

> conversionsAgeMarital$AgeGroup <- as.character(conversionsAgeMarital$AgeGroup)

> conversionsAgeMarital$AgeGroup[is.na(conversionsAgeMarital$AgeGroup)] <- "70+"

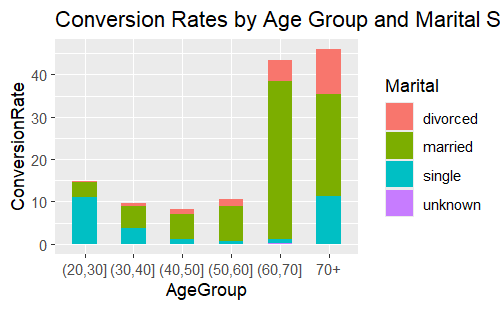
>

> #visualizing conversions by age group and marrital status

> ggplot(conversionsAgeMarital, aes(x=AgeGroup, y=ConversionRate, fill=Marital)) +

+ geom\_bar(width=0.5, stat = "identity") +

+ labs(title="Conversion Rates by Age Group and Marital Status")



```{r}

#group the data

conversionsJob <- data %>%

group\_by(Job=job) %>%

summarize(TotalCount=n(), NumberConversions=sum(y)) %>%

mutate(ConversionRate=NumberConversions/TotalCount\*100) %>%

arrange(desc(ConversionRate))

#order the jobs DESC for the bar chart

conversionsJob$Job <- factor(conversionsJob$Job,

levels = conversionsJob$Job[order(-conversionsJob$ConversionRate)])

# visualizing conversions by job

ggplot(conversionsJob, aes(x=Job, y=ConversionRate)) +

geom\_bar(width=0.5, stat = "identity", fill="darkgreen") +

labs(title="Conversion Rates by Job") +

theme(axis.text.x = element\_text(angle = 90))

```

> #group the data

> conversionsJob <- data %>%

+ group\_by(Job=job) %>%

+ summarize(TotalCount=n(), NumberConversions=sum(y)) %>%

+ mutate(ConversionRate=NumberConversions/TotalCount\*100) %>%

+ arrange(desc(ConversionRate))

>

> #order the jobs DESC for the bar chart

> conversionsJob$Job <- factor(conversionsJob$Job,

+ levels = conversionsJob$Job[order(-conversionsJob$ConversionRate)])

>

> # visualizing conversions by job

> ggplot(conversionsJob, aes(x=Job, y=ConversionRate)) +

+ geom\_bar(width=0.5, stat = "identity", fill="darkgreen") +

+ labs(title="Conversion Rates by Job") +

+ theme(axis.text.x = element\_text(angle = 90))

> #group the data

> conversionsJob <- data %>%

+ group\_by(Job=job) %>%

+ summarize(TotalCount=n(), NumberConversions=sum(y)) %>%

+ mutate(ConversionRate=NumberConversions/TotalCount\*100) %>%

+ arrange(desc(ConversionRate))

>

> #order the jobs DESC for the bar chart

> conversionsJob$Job <- factor(conversionsJob$Job,

+ levels = conversionsJob$Job[order(-conversionsJob$ConversionRate)])

>

> # visualizing conversions by job

> ggplot(conversionsJob, aes(x=Job, y=ConversionRate)) +

+ geom\_bar(width=0.5, stat = "identity", fill="darkgreen") +

+ labs(title="Conversion Rates by Job") +

+ theme(axis.text.x = element\_text(angle = 90))

> #group the data

> conversionsJob <- data %>%

+ group\_by(Job=job) %>%

+ summarize(TotalCount=n(), NumberConversions=sum(y)) %>%

+ mutate(ConversionRate=NumberConversions/TotalCount\*100) %>%

+ arrange(desc(ConversionRate))

>

> #order the jobs DESC for the bar chart

> conversionsJob$Job <- factor(conversionsJob$Job,

+ levels = conversionsJob$Job[order(-conversionsJob$ConversionRate)])

>

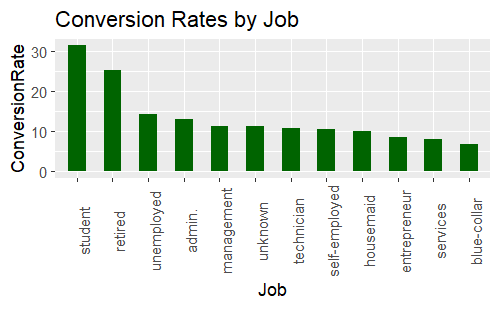
> # visualizing conversions by job

> ggplot(conversionsJob, aes(x=Job, y=ConversionRate)) +

+ geom\_bar(width=0.5, stat = "identity", fill="darkgreen") +

+ labs(title="Conversion Rates by Job") +

+ theme(axis.text.x = element\_text(angle = 90))



## \*\*Conversions by education\*\*

```{r}

#group the data

conversionsEdu <- data %>%

group\_by(Education=education) %>%

summarize(TotalCount=n(), NumberConversions=sum(y)) %>%

mutate(ConversionRate=NumberConversions/TotalCount\*100) %>%

arrange(desc(ConversionRate))

#order DESC for the bar chart

conversionsEdu$Education <- factor(conversionsEdu$Education,

levels = conversionsEdu$Education[order(-conversionsEdu$ConversionRate)])

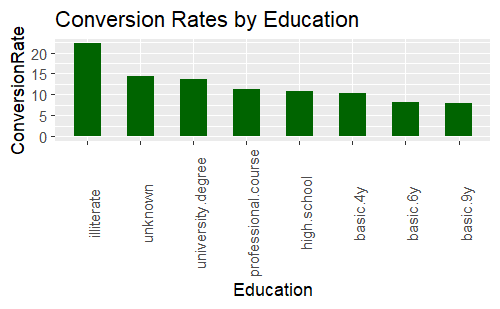
#visualizing conversions by education

ggplot(conversionsEdu, aes(x=Education, y=ConversionRate)) +

geom\_bar(width=0.5, stat = "identity", fill="darkgreen") +

labs(title="Conversion Rates by Education") +

theme(axis.text.x = element\_text(angle = 90)) ```



#group the data

> conversionsEdu <- data %>%

+ group\_by(Education=education) %>%

+ summarize(TotalCount=n(), NumberConversions=sum(y)) %>%

+ mutate(ConversionRate=NumberConversions/TotalCount\*100) %>%

+ arrange(desc(ConversionRate))

>

> #order DESC for the bar chart

> conversionsEdu$Education <- factor(conversionsEdu$Education,

+ levels = conversionsEdu$Education[order(-conversionsEdu$ConversionRate)])

> #visualizing conversions by education

> ggplot(conversionsEdu, aes(x=Education, y=ConversionRate)) +

+ geom\_bar(width=0.5, stat = "identity", fill="darkgreen") +

+ labs(title="Conversion Rates by Education") +

+ theme(axis.text.x = element\_text(angle = 90))

> #group the data

> conversionsEdu <- data %>%

+ group\_by(Education=education) %>%

+ summarize(TotalCount=n(), NumberConversions=sum(y)) %>%

+ mutate(ConversionRate=NumberConversions/TotalCount\*100) %>%

+ arrange(desc(ConversionRate))

>

> #order DESC for the bar chart

> conversionsEdu$Education <- factor(conversionsEdu$Education,

+ levels = conversionsEdu$Education[order(-conversionsEdu$ConversionRate)])

> #visualizing conversions by education

> ggplot(conversionsEdu, aes(x=Education, y=ConversionRate)) +

+ geom\_bar(width=0.5, stat = "identity", fill="darkgreen") +

+ labs(title="Conversion Rates by Education") +

+ theme(axis.text.x = element\_text(angle = 90))

> #group the data

> conversionsEdu <- data %>%

+ group\_by(Education=education) %>%

+ summarize(TotalCount=n(), NumberConversions=sum(y)) %>%

+ mutate(ConversionRate=NumberConversions/TotalCount\*100) %>%

+ arrange(desc(ConversionRate))

>

> #order DESC for the bar chart

> conversionsEdu$Education <- factor(conversionsEdu$Education,

+ levels = conversionsEdu$Education[order(-conversionsEdu$ConversionRate)])

> #visualizing conversions by education

> ggplot(conversionsEdu, aes(x=Education, y=ConversionRate)) +

+ geom\_bar(width=0.5, stat = "identity", fill="darkgreen") +

+ labs(title="Conversion Rates by Education") +

+ theme(axis.text.x = element\_text(angle = 90))

```{r}

#group the data

conversionsDefaultCredit <- data %>%

group\_by(HasCredit=default) %>%

summarize(TotalCount=n(), NumberConversions=sum(y)) %>%

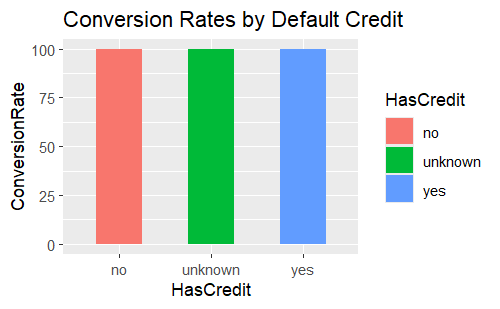
mutate(ConversionRate=NumberConversions/TotalCount\*100) %>%

arrange(desc(ConversionRate))

#visualizing the data

ggplot(conversionsDefaultCredit, aes(x=HasCredit, y=ConversionRate, fill=HasCredit)) + geom\_bar(width0.5, stat = "identity") +

labs(title="Conversion Rates by Default Credit"



#group the data

> conversionsDefaultCredit <- data %>%

+ group\_by(HasCredit=default) %>%

+ summarize(TotalCount=n(), NumberConversions=sum(y)) %>%

+ mutate(ConversionRate=NumberConversions/TotalCount\*100) %>%

+ arrange(desc(ConversionRate))

>

> #visualizing the data

> ggplot(conversionsDefaultCredit, aes(x=HasCredit, y=ConversionRate, fill=HasCredit)) +

+ geom\_bar(width=0.5, stat = "identity") +

+ labs(title="Conversion Rates by Default Credit")

> #group the data - housing loan

> conversionsHousing <- data %>%

+ group\_by(HousingLoan=housing) %>%

+ summarize(TotalCount=n(), NumberConversions=sum(y)) %>%

+ mutate(ConversionRate=NumberConversions/TotalCount\*100) %>%

+ arrange(desc(ConversionRate))

>

> #visualizing the data

> ggplot(conversionsHousing, aes(x=HousingLoan, y=ConversionRate, fill=HousingLoan)) +

+ geom\_bar(width=0.5, stat = "identity") +

+ labs(title="Conversion Rates by Housing Loan")

>

```{r}

#group the data - housing loan

conversionsHousing <- data %>%

group\_by(HousingLoan=housing) %>%

summarize(TotalCount=n(), NumberConversions=sum(y)) %>%

mutate(ConversionRate=NumberConversions/TotalCount\*100) %>%

arrange(desc(ConversionRate))

#visualizing the data

ggplot(conversionsHousing, aes(x=HousingLoan, y=ConversionRate, fill=HousingLoan)) +

geom\_bar(width=0.5, stat = "identity") +

labs(title="Conversion Rates by Housing Loan")

#group the data - personal loan

conversionsLoan <- data %>%

group\_by(Loan=loan) %>%

summarize(TotalCount=n(), NumberConversions=sum(y)) %>%

mutate(ConversionRate=NumberConversions/TotalCount\*100) %>%

arrange(desc(ConversionRate))

#visualizing the data

ggplot(conversionsLoan, aes(x=Loan, y=ConversionRate, fill=Loan)) +

geom\_bar(width=0.5, stat = "identity") +

labs(title="Conversion Rates by Personal Loan")

```

#group the data - housing loan

> conversionsHousing <- data %>%

+ group\_by(HousingLoan=housing) %>%

+ summarize(TotalCount=n(), NumberConversions=sum(y)) %>%

+ mutate(ConversionRate=NumberConversions/TotalCount\*100) %>%

+ arrange(desc(ConversionRate))

>

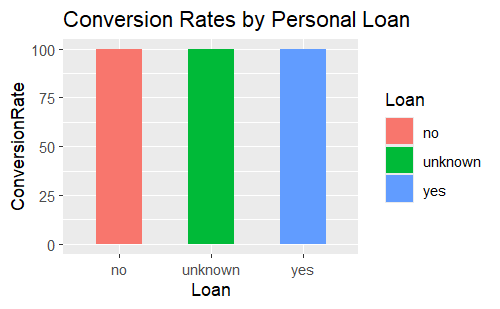
> #visualizing the data

> ggplot(conversionsHousing, aes(x=HousingLoan, y=ConversionRate, fill=HousingLoan)) +

+ geom\_bar(width=0.5, stat = "identity") +

+ labs(title="Conversion Rates by Housing Loan")

>



#group the data - personal loan

conversionsLoan <- data %>%

group\_by(Loan=loan) %>%

summarize(TotalCount=n(), NumberConversions=sum(y)) %>%

mutate(ConversionRate=NumberConversions/TotalCount\*100) %>%

arrange(desc(ConversionRate))

#visualizing the data

ggplot(conversionsLoan, aes(x=Loan, y=ConversionRate, fill=Loan)) +

geom\_bar(width=0.5, stat = "identity") +

labs(title="Conversion Rates by Personal Loan")

```

#group the data - personal loan

> conversionsLoan <- data %>%

+ group\_by(Loan=loan) %>%

+ summarize(TotalCount=n(), NumberConversions=sum(y)) %>%

+ mutate(ConversionRate=NumberConversions/TotalCount\*100) %>%

+ arrange(desc(ConversionRate))

>

> #visualizing the data

> ggplot(conversionsLoan, aes(x=Loan, y=ConversionRate, fill=Loan)) +

+ geom\_bar(width=0.5, stat = "identity") +

+ labs(title="Conversion Rates by Personal Loan")

## \*\*Conversions by contact type\*\*

```{r}

conversionsContact <- data %>%

group\_by(Contact=contact) %>%

summarize(TotalCount=n(), NumberConversions=sum(y)) %>%

mutate(ConversionRate=NumberConversions/TotalCount\*100) %>%

arrange(desc(ConversionRate))

head(conversionsContact)

```

A tibble:2 × 4

|  |
| --- |
|  |

| **Contact**  <chr> | **TotalCount**  <int> |  |
| --- | --- | --- |
| cellular | 26144 |  |
| telephone | 15044 |  |

2 rows | 1-2 of 4 columns

> conversionsContact <- data %>%

+ group\_by(Contact=contact) %>%

+ summarize(TotalCount=n(), NumberConversions=sum(y)) %>%

+ mutate(ConversionRate=NumberConversions/TotalCount\*100) %>%

+ arrange(desc(ConversionRate))

>

> head(conversionsContact)

```{r}

# group the data by months

conversionsMonth <- data %>%

group\_by(Month=month) %>%

summarize(TotalCount=n(), NumberConversions=sum(y)) %>%

mutate(ConversionRate=NumberConversions/TotalCount\*100) %>%

arrange(desc(ConversionRate))

#reorder DESC

conversionsMonth$Month <- factor(conversionsMonth$Month,

levels = conversionsMonth$Month[order(-conversionsMonth$ConversionRate)])

#visualizing the data

ggplot(conversionsMonth, aes(x=Month, y=ConversionRate)) +

geom\_bar(width=0.5, stat = "identity", fill="darkgreen") +

labs(title="Conversion Rates by Last Contact Month") +

theme(axis.text.x = element\_text(angle = 90))

```

# group the data by months

> conversionsMonth <- data %>%

+ group\_by(Month=month) %>%

+ summarize(TotalCount=n(), NumberConversions=sum(y)) %>%

+ mutate(ConversionRate=NumberConversions/TotalCount\*100) %>%

+ arrange(desc(ConversionRate))

>

> #reorder DESC

> conversionsMonth$Month <- factor(conversionsMonth$Month,

+ levels = conversionsMonth$Month[order(-conversionsMonth$ConversionRate)])

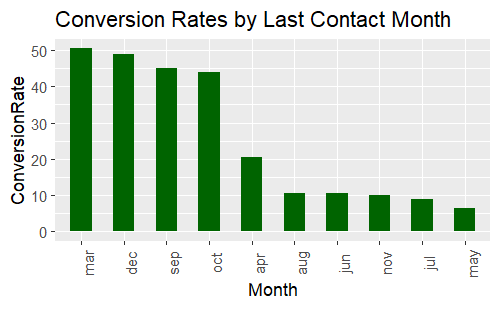
> #visualizing the data

> ggplot(conversionsMonth, aes(x=Month, y=ConversionRate)) +

+ geom\_bar(width=0.5, stat = "identity", fill="darkgreen") +

+ labs(title="Conversion Rates by Last Contact Month") +

+ theme(axis.text.x = element\_text(angle = 90))



## \*\*Conversions by the last contact day of a week\*\*

```{r}

#group the data by days of a week

conversionsDayOfWeek <- data %>%

group\_by(Day\_Of\_Week=day\_of\_week) %>%

summarize(TotalCount=n(), NumberConversions=sum(y)) %>%

mutate(ConversionRate=NumberConversions/TotalCount\*100) %>%

arrange(desc(ConversionRate))

#reorder DESC

conversionsDayOfWeek$Day\_Of\_Week <- factor(conversionsDayOfWeek$Day\_Of\_Week,

levels = c("mon", "tue", "wed", "thu", "fri"))

#visualizing the data

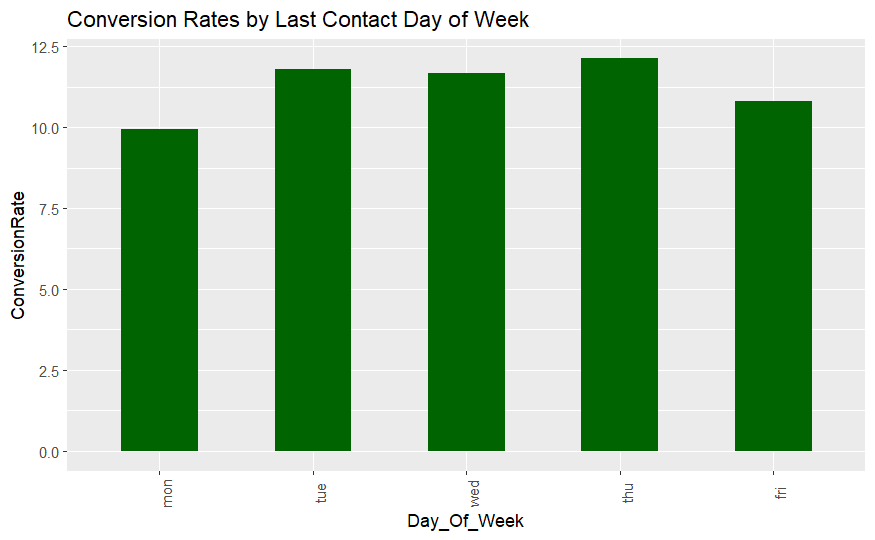
ggplot(conversionsDayOfWeek, aes(x=Day\_Of\_Week, y=ConversionRate)) +

geom\_bar(width=0.5, stat = "identity", fill="darkgreen") +

labs(title="Conversion Rates by Last Contact Day of Week") +

theme(axis.text.x = element\_text(angle = 90))

```



#group the data by days of a week

> conversionsDayOfWeek <- data %>%

+ group\_by(Day\_Of\_Week=day\_of\_week) %>%

+ summarize(TotalCount=n(), NumberConversions=sum(y)) %>%

+ mutate(ConversionRate=NumberConversions/TotalCount\*100) %>%

+ arrange(desc(ConversionRate))

>

> #reorder DESC

> conversionsDayOfWeek$Day\_Of\_Week <- factor(conversionsDayOfWeek$Day\_Of\_Week,

+ levels = c("mon", "tue", "wed", "thu", "fri"))

> #visualizing the data

> ggplot(conversionsDayOfWeek, aes(x=Day\_Of\_Week, y=ConversionRate)) +

+ geom\_bar(width=0.5, stat = "identity", fill="darkgreen") +

+ labs(title="Conversion Rates by Last Contact Day of Week") +

+ theme(axis.text.x = element\_text(angle = 90))

## \*\*Correlation between subscribing to a term deposit and call duration\*\*

```{r}

data\_duration <- data %>%

group\_by(Subscribed=y) %>%

summarise(Average\_Duration=mean(duration))

head(data\_duration)

```

A tibble:2 × 2

|  |
| --- |
|  |

| **Subscribed**  <int> | **Average\_Duration**  <dbl> |
| --- | --- |
| 0 | 220.8448 |
| 1 | 553.1912 |

2 rows

data\_duration <- data %>%

+ group\_by(Subscribed=y) %>%

+ summarise(Average\_Duration=mean(duration))

> head(data\_duration)

## \*\*Conversions by the number of contacts performed during the campaign\*\*

```{r}

conversionsCamp <- data %>%

group\_by(Campaign=campaign) %>%

summarize(TotalCount=n(), NumberConversions=sum(y)) %>%

mutate(ConversionRate=NumberConversions/TotalCount\*100) %>%

arrange(desc(ConversionRate))

head(conversionsCamp)

```

A tibble:6 × 4

|  |
| --- |
|  |

| **Campaign**  <int> | **TotalCount**  <int> |  |
| --- | --- | --- |
| 1 | 17642 |  |
| 2 | 10570 |  |
| 3 | 5341 |  |
| 4 | 2651 |  |
| 6 | 979 |  |
| 5 | 1599 |  |

6 rows | 1-2 of 4 columns

conversionsCamp <- data %>%

+ group\_by(Campaign=campaign) %>%

+ summarize(TotalCount=n(), NumberConversions=sum(y)) %>%

+ mutate(ConversionRate=NumberConversions/TotalCount\*100) %>%

+ arrange(desc(ConversionRate))

>

> head(conversionsCamp)

## \*\*Conversions by the outcome of the previous campaign\*\*

```{r}

#group the data by the previous outcome

conversionsPOutcome <- data %>%

group\_by(Previous\_Outcome=poutcome) %>%

summarize(TotalCount=n(), NumberConversions=sum(y)) %>%

mutate(ConversionRate=NumberConversions/TotalCount\*100) %>%

arrange(desc(ConversionRate))

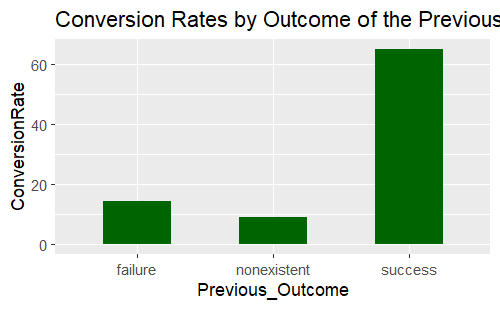
# visualizing the data

ggplot(conversionsPOutcome, aes(x=Previous\_Outcome, y=ConversionRate)) +

geom\_bar(width=0.5, stat = "identity", fill="darkgreen") +

labs(title="Conversion Rates by Outcome of the Previous Campaign")

```



|  |
| --- |
| #group the data by the previous outcome  > conversionsPOutcome <- data %>%  + group\_by(Previous\_Outcome=poutcome) %>%  + summarize(TotalCount=n(), NumberConversions=sum(y)) %>%  + mutate(ConversionRate=NumberConversions/TotalCount\*100) %>%  + arrange(desc(ConversionRate))  >  > # visualizing the data  > ggplot(conversionsPOutcome, aes(x=Previous\_Outcome, y=ConversionRate)) +  + geom\_bar(width=0.5, stat = "identity", fill="darkgreen") +  + labs(title="Conversion Rates by Outcome of the Previous Campaign") |
|  |
| |  | | --- | | > | |