

Import:

```
> library(readxl)
>
> data <- read_excel("/Users/edengreenfest/Downloads/Typical_Employee_Survey_Data.xlsx", col_names = FALSE)
New names:
• `` -> `...1`
• `` -> `...2`
• `` -> `...3`
• `` -> `...4`
• `` -> `...5`
• `` -> `...6`
• `` -> `...7`
• `` -> `...8`
• `` -> `...9`
• `` -> `...10`
• `` -> `...11`
> file_ext <- tools::file_ext("/Users/edengreenfest/Downloads/Typical_Employee_Survey_Data.xlsx")
> print(file_ext)
[1] "xlsx"
> colnames(data) <- c("Age", "Gender", "Job Satisfaction", "Job Characteristics", "Years Worked for Present Employer", "Promotion Likeliness", "Can Make Decisions", "Budgetary Decisions", "Pride in Organization", "Loyalty", "Culture")
>
**Later went back and updated: > Typical_Employee_Survey_Data <-
read_excel("Downloads/Typical_Employee_Survey_Data.xlsx",
+          col_names = TRUE)
```

Columns:

```
> colnames(data)
[1] "Age"                "Gender"
[3] "Job Satisfaction"   "Job Characteristics"
[5] "Years Worked for Present Employer" "Promotion Likeliness"
[7] "Can Make Decisions" "Budgetary Decisions"
[9] "Pride in Organization" "Loyalty"
[11] "Culture"
>
```

Later went back and fixed:

```
colnames(Typical_Employee_Survey_Data) <- c("Age", "Gender", "Job Satisfaction",
+      "Job Characteristics",
+      "Years Worked for Present Employer",
+      "Promotion Likeliness",
+      "Can Make Decisions",
+      "Budgetary Decisions",
+      "Pride in Organization",
+      "Loyalty", "Culture")
```

This step gives a descriptive name to each column, making the dataset more readable and accessible.

5 point summary: The 5-point summary and boxplot helps give an overview of the numerical data and visualize its distribution.

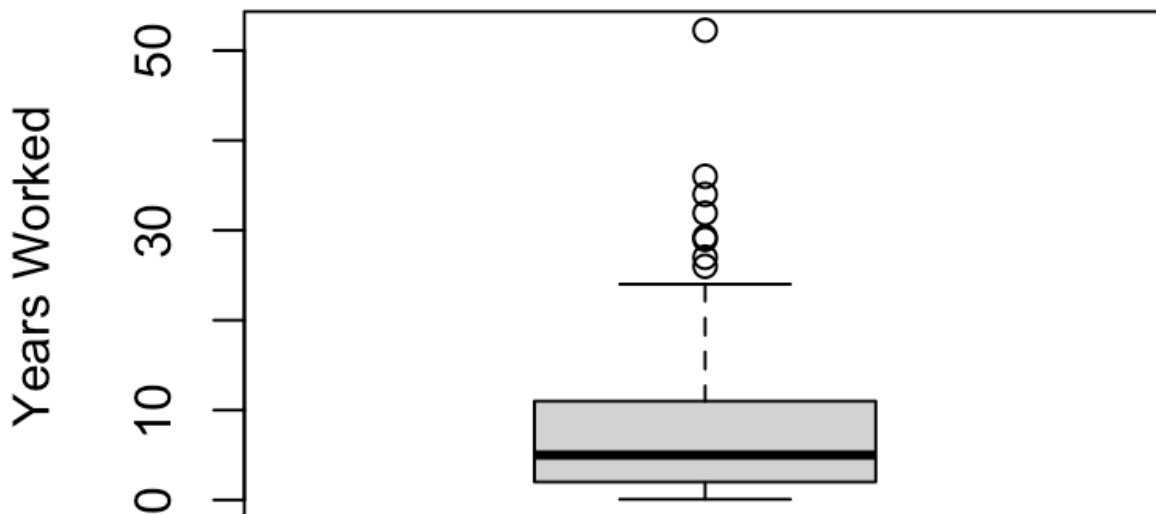
```
> summary(data)
```

Age	Gender	Job Satisfaction	Job Characteristics
Min. :20.0	Min. :1.000	Min. :1.000	Min. :1.000
1st Qu.:33.0	1st Qu.:1.000	1st Qu.:1.000	1st Qu.:2.000
Median :38.0	Median :1.000	Median :1.000	Median :5.000
Mean :39.1	Mean :1.426	Mean :1.631	Mean :3.705
3rd Qu.:44.0	3rd Qu.:2.000	3rd Qu.:2.000	3rd Qu.:5.000
Max. :64.0	Max. :2.000	Max. :4.000	Max. :5.000

Years Worked for Present Employer	Promotion Likliness	Can Make Decisions	Budgetary Decisions
Min. : 0.080	Min. :1.000	Min. :1.000	Min. :1.000
1st Qu.: 2.000	1st Qu.:2.000	1st Qu.:2.000	1st Qu.:1.000
Median : 5.000	Median :4.000	Median :2.000	Median :1.000
Mean : 8.224	Mean :3.287	Mean :2.279	Mean :1.451
3rd Qu.:11.000	3rd Qu.:5.000	3rd Qu.:3.000	3rd Qu.:2.000
Max. :52.250	Max. :5.000	Max. :4.000	Max. :2.000

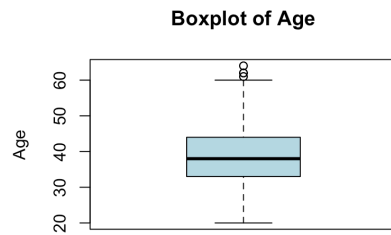
Pride in Organization	Loyalty	Culture
Min. :1.000	Min. :1.000	Min. :1.000
1st Qu.:1.000	1st Qu.:2.000	1st Qu.:1.000
Median :2.000	Median :4.000	Median :2.000
Mean :1.787	Mean :3.254	Mean :1.787
3rd Qu.:2.000	3rd Qu.:4.000	3rd Qu.:2.000
Max. :4.000	Max. :5.000	Max. :4.000

Boxplot for Years Worked

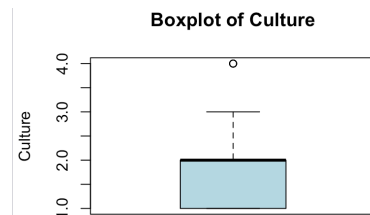
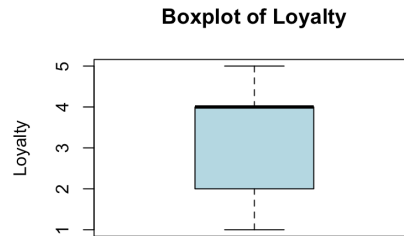
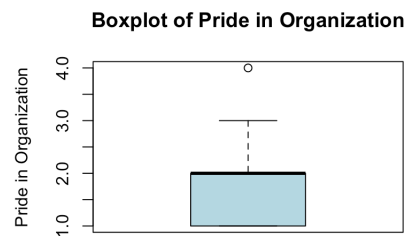
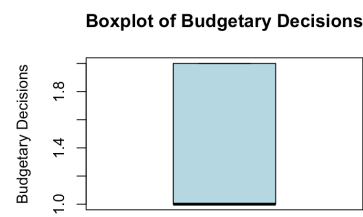
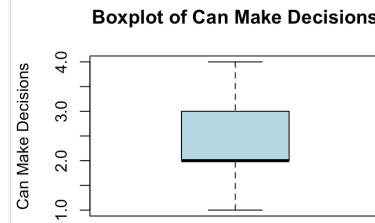
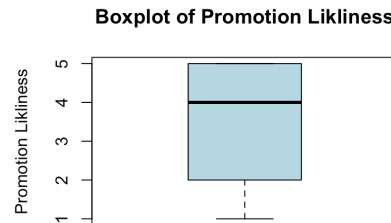


boxplot for each column:

```
> boxplot(data$Age, main = "Boxplot of Age", ylab = "Age", col = "lightblue", border = "black")
```



```
boxplot(data$`Promotion Likliness`, main = "Boxplot of Promotion Likliness", ylab = "Promotion Likliness", col = "lightblue", border = "black")
```



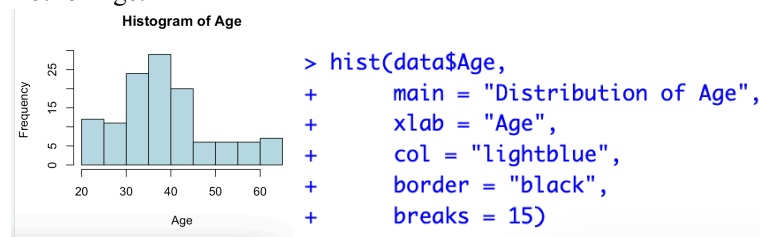
Converting column to factor: Some columns are categorical so they should be converted to factors to make analysis easier.

```
data$Gender <- factor(data$Gender, levels = c(1, 2), labels = c("Male", "Female"))
data$`Job Satisfaction` <- factor(data$`Job Satisfaction`, levels = c(1, 2, 3, 4), labels = c("Very Satisfied",
"Moderately Satisfied", "A Little Dissatisfied", "Very Dissatisfied"))
data$`Job Characteristics` <- factor(data$`Job Characteristics`, levels = c(1, 2, 3, 4, 5), labels = c("High Income",
"No Danger of Being Fired", "Flexible Hours", "Opportunities for Advancement", "Enjoying the Work"))
data$`Promotion Likliness` <- factor(data$`Promotion Likliness`, levels = c(1, 2, 3, 4, 5), labels = c("Very Likely",
"Likely", "Not Sure", "Unlikely", "Very Unlikely"))
data$`Can Make Decisions` <- factor(data$`Can Make Decisions`, levels = c(1, 2, 3, 4), labels = c("Always", "Much
of the Time", "Sometimes", "Never"))
data$`Budgetary Decisions` <- factor(data$`Budgetary Decisions`, levels = c(1, 2), labels = c("Yes", "No"))
data$`Pride in Organization` <- factor(data$`Pride in Organization`, levels = c(1, 2, 3, 4), labels = c("Very Proud",
"Somewhat Proud", "Indifferent", "Not At All Proud"))
data$Loyalty <- factor(data$Loyalty, levels = c(1, 2, 3, 4, 5), labels = c("Very Likely", "Likely", "Not Sure",
"Unlikely", "Very Unlikely"))
data$Culture <- factor(data$Culture, levels = c(1, 2, 3, 4, 5), labels = c("Very Good", "Good", "So So", "Bad", "Very
Bad"))
```

```
> str(data)
tibble [122 × 11] (S3: tbl_df/tbl/data.frame)
 $ Age                : num [1:122] 35 33 23 60 35 34 61 59 37 30 ...
 $ Gender              : Factor w/ 2 levels "Male","Female": NA NA NA NA NA |
 ...
 $ Job Satisfaction    : Factor w/ 4 levels "Very Satisfied",...: NA NA NA NA
 NA ...
 $ Job Characteristics : Factor w/ 5 levels "High Income",...: NA NA NA NA NA
 ...
 $ Years Worked for Present Employer: num [1:122] 3 9 1.5 20 3 6 0.75 1.5 3 5 ...
 $ Promotion Likiness  : Factor w/ 5 levels "Very Likely",...: NA NA NA NA NA
 ...
 $ Can Make Decisions  : Factor w/ 4 levels "Always","Much of the Time",...: |
 A NA NA NA NA ...
 $ Budgetary Decisions : Factor w/ 2 levels "Yes","No": NA NA NA NA NA NA
 $ Pride in Organization : Factor w/ 4 levels "Very Proud","Somewhat Proud",..
 NA NA NA NA NA ...
 $ Loyalty             : Factor w/ 5 levels "Very Likely",...: NA NA NA NA NA
 ...
 $ Culture             : Factor w/ 5 levels "Very Good","Good",...: NA NA NA |
 A NA ...
>
```

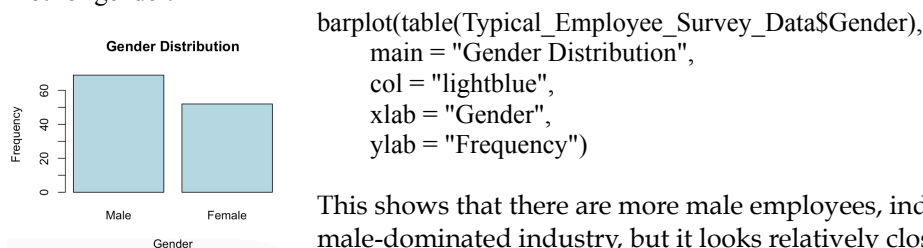
Plots: By plotting each column individually, we can create useful visualizations based on their data type.

Plot for Age:



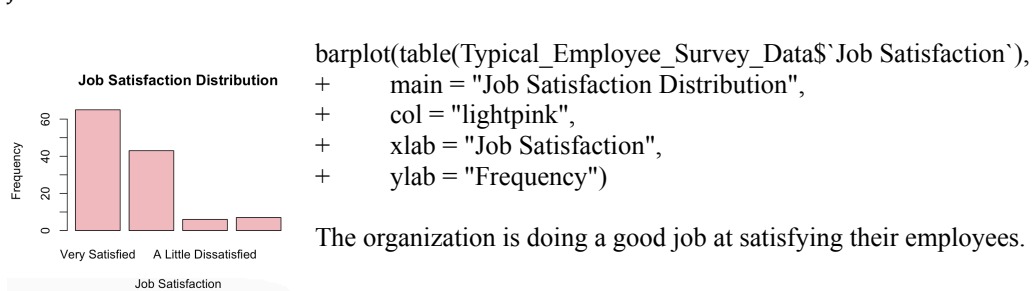
The histogram of age shows the distribution of employees' ages. We can see most employees are clustered around a particular age range (35-45 ish) which suggests that the company might have middle aged workforce.

Plot for gender:



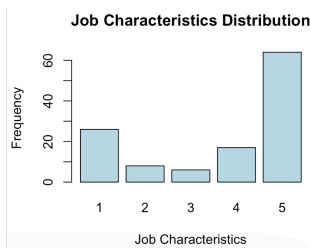
This shows that there are more male employees, indicating a slightly more male-dominated industry, but it looks relatively close.

Job Satisfaction



The organization is doing a good job at satisfying their employees.

Job Characteristics



```
barplot(table(Typical_Employee_Survey_Data$`Job Characteristics`),
main = "Job Characteristics Distribution",
col = "lightblue",
xlab = "Job Characteristics",
ylab = "Frequency")
```

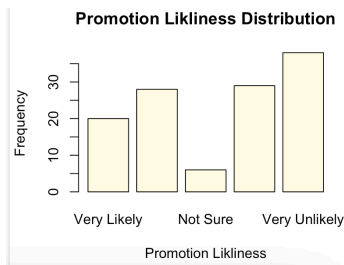
Based on the plot it looks like characteristic 1 and 5 are most common amongst employees.

Years Worked for Present Employer:

```
hist(data$`Years Worked for Present Employer`,
main = "Years Worked for Present Employer",
xlab = "Years Worked",
col = "lightyellow",
border = "black",
breaks = 10)
```



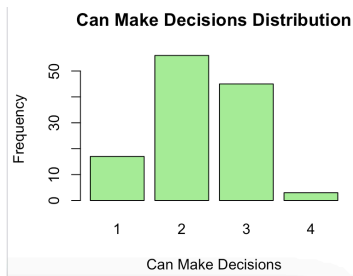
Most employees seem to be relatively new to the organization, starting with the past ten years or so.



```
Promotion Likelihood:
barplot(table(Typical_Employee_Survey_Data$`Promotion Likliness`),
main = "Promotion Likliness Distribution",
col = "lightyellow",
xlab = "Promotion Likliness",
ylab = "Frequency")
```

A lot of employees seem to have strong opinions about whether or not they are likely to get promoted.

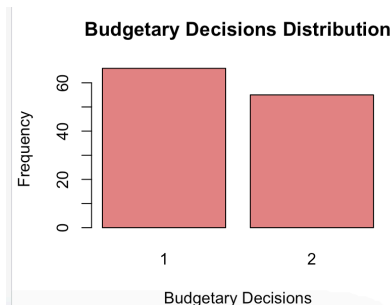
Can Make Decisions:



```
barplot(table(Typical_Employee_Survey_Data$`Can Make Decisions`),
main = "Can Make Decisions Distribution",
col = "lightgreen",
xlab = "Can Make Decisions",
ylab = "Frequency")
```

Majority of employees seem to be between a level 2 and 3 when it comes to being able to make decisions that impact their tasks.

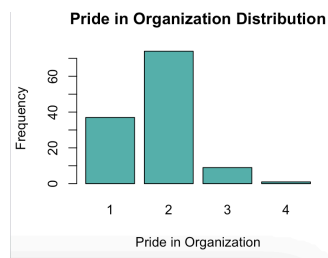
Budgetary Decisions:



```
barplot(table(Typical_Employee_Survey_Data$`Budgetary Decisions`),
main = "Budgetary Decisions Distribution",
col = "lightcoral",
xlab = "Budgetary Decisions",
ylab = "Frequency")
```

The split looks pretty even on whether employees can make budgetary decisions.

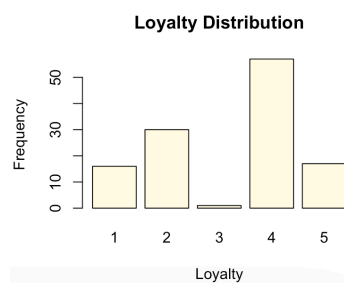
Pride in Organization:



```
barplot(table(Typical_Employee_Survey_Data$`Pride in Organization`),
  main = "Pride in Organization Distribution",
  col = "lightseagreen",
  xlab = "Pride in Organization",
  ylab = "Frequency")
```

Many of the employees seem to have a higher level of pride in working for this organization.

Loyalty:



```
barplot(table(Typical_Employee_Survey_Data$`Loyalty`),
  main = "Loyalty Distribution",
  col = "lightyellow",
  xlab = "Loyalty",
  ylab = "Frequency")
```

It seems like it's a pretty even split about whether employees would accept another job when given a higher salary. Interesting to note that there seems to be a clear yes vs no, rather than unsure.

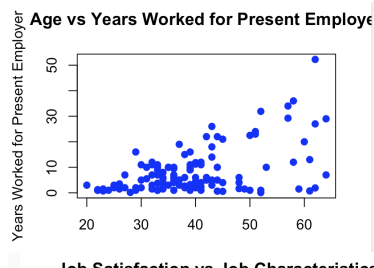
Culture:



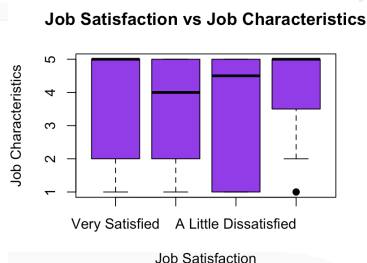
```
barplot(table(Typical_Employee_Survey_Data$Culture),
  main = "Culture Distribution",
  col = "lightcoral",
  xlab = "Culture",
  ylab = "Frequency")
```

Relationships:

Age vs. Job Characteristic



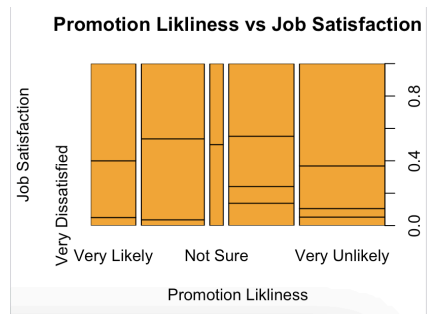
```
plot(Typical_Employee_Survey_Data$Age,
  + Typical_Employee_Survey_Data$`Years Worked for Present Employer`,
  + main = "Age vs Years Worked for Present Employer",
  + xlab = "Age",
  + ylab = "Years Worked for Present Employer",
  + pch = 19,
  + col = "blue")
```



Job Satisfaction vs Job Characteristics

```
plot(Typical_Employee_Survey_Data$`Job Satisfaction`,
  Typical_Employee_Survey_Data$`Job Characteristics`,
  main = "Job Satisfaction vs Job Characteristics",
```

```
xlab = "Job Satisfaction",
ylab = "Job Characteristics",
pch = 19,
col = "purple")
```



Promotion Likliness vs Job Satisfaction

```
> plot(Typical_Employee_Survey_Data$`Promotion Likliness`,
+      Typical_Employee_Survey_Data$`Job Satisfaction`,
+      main = "Promotion Likliness vs Job Satisfaction",
+      xlab = "Promotion Likliness",
+      ylab = "Job Satisfaction",
+      pch = 19,
+      col = "orange")
```

Analysis:

Age vs. Years Worked: Older employees tend to stay longer with their employersthis shows job stability over time.

Job Satisfaction vs. Job Characteristics: Stronger job characteristics (good work environment, roles, etc.) seems to result in higher job satisfaction.

Promotion Likelihood vs. Job Satisfaction: Employees who perceive a higher chance of promotion tend to be more satisfied with their jobs.