



Herzlichen Glückwunsch! Sie haben bestanden!

ZUM BESTEHEN 80 % oder höher

Lernen Sie weiter

BEWERTUNG
84,61 %

Course challenge

NEUESTE EINREICHUNGSBEWERTUNG

84.61%

1. Scenario 1, questions 1-7

1 / 1 Punkt

As part of the data science team at Gourmet Analytics, you use data analytics to advise companies in the food industry. You clean, organize, and visualize data to arrive at insights that will benefit your clients. As a member of a collaborative team, sharing your analysis with others is an important part of your job.

Your current client is Chocolate and Tea, an up-and-coming chain of cafes.



The eatery combines an extensive menu of fine teas with chocolate bars from around the world. Their diverse selection includes everything from plantain milk chocolate, to tangerine white chocolate, to dark chocolate with pistachio and fig. The encyclopedic list of chocolate bars is the basis of Chocolate and Tea's brand appeal. Chocolate bar sales are the main driver of revenue.

Chocolate and Tea aims to serve chocolate bars that are highly rated by professional critics. They also continually adjust the menu to make sure it reflects the global diversity of chocolate production. The management team regularly updates the chocolate bar list in order to align with the latest ratings and to ensure that the list contains bars from a variety of countries.

They've asked you to collect and analyze data on the latest chocolate ratings. In particular, they'd like to know which countries produce the highest-rated bars of super dark chocolate (a high percentage of cocoa). This data will help them create their next chocolate bar menu.

Your team has received a dataset that features the latest ratings for thousands of chocolates from around the world. Click [here](#) to access the dataset. Given the data and the nature of the work you will do for your client, your team agrees to use R for this project.

Your supervisor asks you to write a short summary of the benefits of using R for the project. Which of the following benefits would you include in your summary? Select all that apply.

☒ Easily reproduce and share the analysis

✓ Richtig

The benefits of using R for the project include the ability to quickly process lots of data and create high-quality data visualizations. You can also easily reproduce and share your analysis.

☒ Create high-quality data visualizations

✓ Richtig

The benefits of using R for the project include the ability to quickly process lots of data and create high-quality data visualizations. You can also easily reproduce and share your analysis.

☐ Define a problem and ask the right questions

☒ Quickly process lots of data

✓ Richtig

The benefits of using R for the project include the ability to quickly process lots of data and create high-quality data visualizations. You can also easily reproduce and share your analysis.

2. Scenario 1, continued

1 / 1 Punkt

Before you begin working with your data, you need to import it and save it as a data frame. To get started, you open your RStudio workspace and load all the necessary libraries and packages. You upload a .csv file containing the data to RStudio and store it in a project folder named `flavors_of_cacao.csv`.

You use the `read_csv()` function to import the data from the .csv file. Assume that the name of the data frame is `flavors_df` and the .csv file is in the working directory. What code chunk lets you create the data frame?

- ☐ `flavors_df = read_csv("flavors_of_cacao.csv")`
- ☐ `read_csv("flavors_of_cacao.csv") <- flavors_df`
- ☐ `read_csv(flavors_df <- "flavors_of_cacao.csv")`
- ☒ `flavors_df <- read_csv("flavors_of_cacao.csv")`

✓ Richtig

You write the code chunk: `flavors_df <- read_csv("flavors_of_cacao.csv")`.

In this code chunk:

- `flavors_df` is the name of the data frame that will store the data.
- `<-` is the assignment operator to assign values to the data frame.
- `read_csv()` is the function that will import the data to the data frame.
- `"flavors_of_cacao.csv"` is the file name that `read_csv()` function takes for its argument.

3. Scenario 1, continued

1 / 1 Punkt

Now that you've created a data frame, you want to find out more about how the data is organized. The data frame has hundreds of rows and lots of columns.

Assume the name of your data frame is `flavors_df`. What code chunk lets you review the structure of the data frame?

- ☐ `filter(flavors_df)`
- ☐ `select(flavors_df)`
- ☒ `str(flavors_df)`
- ☐ `summarise(flavors_df)`

✓ Richtig

You write the code chunk `str(flavors_df)`. In this code chunk:

- `str()` is the function that will return the structure of the data frame, and give you high-level information like the column names and the type of data contained in those columns.
- `flavors_df` is the name of the data frame that the `str()` function takes for its argument.

4. Scenario 1, continued

1 / 1 Punkt

Next, you begin to clean your data. When you check out the column headings in your data frame you notice that the first column is named `Company...Maker.if.known`. (Note: The period after *known* is part of the variable name.) For the sake of clarity and consistency, you decide to rename this column `Maker` (without a period at the end).

Assume the first part of your code chunk is:

```
flavors_df %>%
```

What code chunk do you add to change the column name?

- ☐ `rename(Maker %<% Company...Maker.if.known.)`
- ☐ `rename(Company...Maker.if.known. = Maker)`
- ☒ `rename(Maker = Company...Maker.if.known.)`
- ☐ `rename(Company...Maker.if.known %<% Maker)`

✓ Richtig

You write the code chunk `rename(Maker = Company...Maker.if.known.)`.

In this code chunk:

- `rename()` is the function that will change the name of your column.
- Inside the parentheses of the function, write the new name (`Maker`), then an equal sign, then the name you want to change (`Company...Maker.if.known.`).

5. Scenario 1, continued

1 / 1 Punkt

After previewing and cleaning your data, you determine what variables are most relevant to your analysis. Your main focus is on `Rating`, `Cocoa.Percent`, and `Bean.Type`. You decide to use a function to create a new data frame with only these three variables.

Assume the first part of your code chunk is:

```
trimmed_flavors_df <- flavors_df %>%
```

What code chunk do you add to choose the three variables?

- ☐ `filter(Rating + Cocoa.Percent + Bean.Type)`
- ☒ `select(Rating, Cocoa.Percent, Bean.Type)`
- ☐ `summarise(Rating & Cocoa.Percent & Bean.Type)`
- ☐ `separate(Rating, Cocoa.Percent, Bean.Type)`

✓ Richtig

You write the code chunk `select(Rating, Cocoa.Percent, Bean.Type)`. In this code chunk:

- `select()` is the function that lets you select specific variables for your new data frame.
- `select()` takes the names of the variables you want to choose as its argument: `Rating`, `Cocoa.Percent`, `Bean.Type`.

6. Scenario 1, continued

1 / 1 Punkt

Next, you select the basic statistics that can help your team better understand the ratings system in your data.

Assume the first part of your code chunk is:

```
trimmed_flavors_df %>%
```

What code chunk do you add to determine the mean rating for your data?

- ☐ `summarize %>% mean(Rating))`
- ☐ `arrange <- mean(Rating)`
- ☒ `summarize(mean(Rating))`
- ☐ `arrange(mean, Rating)`

✓ Richtig

You write the code chunk `summarize(mean(Rating))`. In this code chunk:

- `summarize()` is the function that lets you display summary statistics.
- In this case, you calculate the mean statistic for the variable `Rating`.

7. Scenario 1, continued

1 / 1 Punkt

After completing your analysis of the rating system, you determine that any rating equal to or greater than 3.5 can be considered a high rating. You also know that Chocolate and Tea considers any bar that contains at least 70% cocoa to be super dark chocolate. You decide to use code to find out which chocolate bars meet these two conditions.

Assume the first part of your code chunk is:

```
best_trimmed_flavors <- trimmed_flavors_df %>%
```

What code chunk do you add to filter the data frame for chocolate bars that contain at least 70% cocoa and have a rating of at least 3.5 points?

- ☐ `filter(Cocoa.Percent > 70, Rating > 3.5)`
- ☐ `filter(Cocoa.Percent >= 70, Rating == 3.5)`
- ☒ `filter(Cocoa.Percent >= 70, Rating >= 3.5)`
- ☐ `filter(Cocoa.Percent == 70, Rating > 3.5)`

✓ Richtig

You write the code chunk `filter(Cocoa.Percent >= 70, Rating >= 3.5)`.

In this code chunk:

- `filter()` is the function that lets you filter your data frame based on specific criteria.
- `Cocoa.Percent` and `Rating` refer to the variables you want to filter.
- The `>=` operator signifies "greater than or equal to."
- The new data frame will show all the values of `Cocoa.Percent` greater than or equal to 70, and all the values of `Rating` greater than or equal to 3.5.

8. Scenario 2, questions 8-13

0 / 1 Punkt

Now that you've cleaned and organized your data, you're ready to create some useful data visualizations. Your team assigns you the task of creating a series of visualizations based on requests from the Chocolate and Tea management team. You decide to use `ggplot2` to create your visuals.

Assume the first part of your code chunk is:

```
ggplot(data = best_trimmed_flavors_df) +
```

What code chunk do you add to the second line to create a bar chart with the variable *Company* on the x-axis?

- ☒ `geom_bar(aes(x = Company))`
- ☐ `geom_bar(mapping = aes(Company))`
- ☐ `geom_bar(mapping = aes(x = Company))`
- ☐ `geom_bar(mapping = aes(x + Company))`

✗ Falsch

Review the section on geoms for a refresher.

9. Your bar chart reveals the locations that produce the highest-rated chocolate bars. To get a better idea of the specific rating for each location, you'd like to highlight each bar.

1 / 1 Punkt

Assume that you are working with the code chunk:

```
ggplot(data = best_trimmed_flavors_df) +  
  geom_bar(mapping = aes(x = Company.Location))
```

How do you change the second line of code to give each bar a different degree of transparency?

- ☐ `geom_bar(mapping = aes(x = Company.Location, fill = Rating))`
- ☐ `geom_bar(mapping = aes(x = Company.Location, color = Rating))`
- ☒ `geom_bar(mapping = aes(x = Company.Location, alpha = Rating))`
- ☐ `geom_bar(mapping = aes(x = Company.Location, y = Rating))`

✓ Richtig

You change the second line of code to `geom_bar(mapping = aes(x = Company.Location, alpha = Rating))` to give each bar a different degree of transparency. In this code chunk:

- Inside the parentheses of the `aes()` function, add a comma after `x = Company.Location` to add a new aesthetic attribute, then write `alpha = Rating` to map the aesthetic `alpha` to the variable `Rating`.
- The specific rating of each location will appear as a different degree of transparency in each bar of your bar chart.

10. Scenario 2, continued

1 / 1 Punkt

A teammate creates a new plot based on the chocolate bar data. The teammate asks you to make some revisions to their code.

Assume your teammate shares the following code chunk:

```
ggplot(data = best_trimmed_flavors_df) +  
  geom_bar(mapping = aes(x = Rating)) +
```

What code chunk do you add to the third line to create wrap around facets of the variable *Rating*?

- ☐ `facet(~Rating)`
- ☐ `facet_wrap(Rating)`
- ☐ `facet_wrap(Rating~)`
- ☒ `facet_wrap(~Rating)`

✓ Richtig

You write the code chunk `facet_wrap(~Rating)`. In this code chunk:

• `facet_wrap()` is the function that lets you create wrap around facets of a variable.

- `facet_wrap()` is the function that lets you create wrap around facets of a variable.
- Inside the parentheses of the `facet_wrap()` function, type a tilde symbol (~) followed by the name of the variable (`Rating`).

11. Scenario 2, continued

0 / 1 Punkt

Your team has created some basic visualizations to explore different aspects of the chocolate bar data. You've volunteered to add titles to the plots. You begin with a scatterplot.

Assume the first part of your code chunk is:

```
ggplot(data = trimmed_flavors_df) +  
  
  geom_point(mapping = aes(x = Cocoa.Percent, y = Rating)) +
```

What code chunk do you add to the third line to add the title *Recommended Bars* to your plot?

- ☐ `labs(title + "Recommended Bars")`
- ☒ `labs(title = Recommended Bars)`
- ☐ `labs(title = "Recommended Bars")`
- ☐ `labs("Recommended Bars")`

✗ Falsch

Review the section on labels and annotations for a refresher.

12. Scenario 2, continued

1 / 1 Punkt

Next, you create a new scatterplot to explore the relationship between different variables. You want to save your plot so you can access it later on. You know that the `ggsave()` function defaults to saving the last plot that you displayed in RStudio, so you're ready to write the code to save your scatterplot.

Assume your first two lines of code are:

```
ggplot(data = trimmed_flavors_df) +  
  
  geom_point(mapping = aes(x = Cocoa.Percent, y = Rating))
```

What code chunk do you add to the third line to save your plot as a png file with *chocolate* as the file name?

- ☐ `ggsave("chocolate")`
- ☐ `ggsave(chocolate.png)`
- ☐ `ggsave("png.chocolate")`
- ☒ `ggsave("chocolate.png")`

✓ Richtig

You write the code chunk `ggsave("chocolate.png")`. In this code chunk:

- Inside the parentheses of the `ggsave()` function, type a quotation mark followed by the file name (`chocolate`), then a period, then the type of file format (`png`), then a closing quotation mark.

13. Scenario 2, continued

1 / 1 Punkt

As a final step in the analysis process, you create a report to document and share your work. Before you share your work with the management team at Chocolate and Tea, you are going to meet with your team and get feedback. Your team wants the documentation to include all your code and display all your visualizations.

You decide to create an R Markdown notebook to document your work. What are your reasons for choosing an R Markdown notebook? Select all that apply.

- ☒ It allows users to run your code

✓ Richtig

You choose an R Markdown notebook to document your work because it lets you record and share every step of your analysis. The notebook allows users to run your code and also displays your data visualizations.

- ☐ It automatically creates a website to show your work

- ☒ It displays your data visualizations

✓ Richtig

You choose an R Markdown notebook to document your work because it lets you record and share every step of your analysis. The notebook allows users to run your code and also displays your data visualizations.

- ☒ It lets you record and share every step of your analysis

✓ Richtig

You choose an R Markdown notebook to document your work because it lets you record and share every step of your analysis. The notebook allows users to run your code and also displays your data visualizations.