



## Integrated Project

Congratulations! You’ve officially completed the first part of the course on the interactive platform. Now it's time to bring together all you've learned so far in your first integrated project, a real-life analytical case study.

When you finish the project, send your work to the project reviewer for assessment. They’ll give you feedback within 24 hours. Use the feedback to make changes, then send the new version back to the project reviewer.

You might get further feedback on the new version. This is completely normal. It’s not uncommon to go through several cycles of feedback and revision.

Your project will be considered complete once the project reviewer approves it.

## Project description

You work for the online store Ice, which sells video games all over the world. User and expert reviews, genres, platforms (e.g. Xbox or PlayStation), and historical data on game sales are available from open sources. You need to identify patterns that determine whether a game succeeds or not. This will allow you to spot potential big winners and plan advertising campaigns.

In front of you is data going back to 2016. Let’s imagine that it’s December 2016 and you’re planning a campaign for 2017.

(The important thing is to get experience working with data. It doesn't really matter whether you're forecasting 2017 sales based on data from 2016 or 2017 sales based on data from 2016.)

The dataset contains the abbreviation ESRB. The Entertainment Software Rating Board evaluates a game's content and assigns an age rating such as Teen or Mature.

## Instructions for completing the project

### Step 1. Open the data file and study the general information

File path:

*/datasets/games.csv* . [Download dataset](#)

### Step 2. Prepare the data

- Replace the column names (make them lowercase).
- Convert the data to the required types.
- Describe the columns where the data types have been changed and why.
- If necessary, decide how to deal with missing values:
  - Explain why you filled in the missing values as you did or why you decided to leave them blank.
  - Why do you think the values are missing? Give possible reasons.
  - Pay attention to the abbreviation TBD (to be determined). Specify how you intend to handle such cases.
- Calculate the total sales (the sum of sales in all regions) for each game and put these values in a separate column.

### Step 3. Analyze the data

- Look at how many games were released in different years. Is the data for every period significant?

