DATA U
CREATING TOMORROW'S DATA LEADERS, TODAY

PROJECT - HS 01

EXCEL PROJECT

DATAJ
CREATING TOMORROWS DATA LEADERS, TODAY

Project summary

The goal of this project is to give you a chance to practically apply the information you received in all 4 weeks of training in excel and receive peer and HS team review on this. You will be creating an Excel dashboard for your business analysis project. This gives you the opportunity to choose a business case and questions that are of particular interest to you. As a result, finding datasets and conducting analysis will be more complex and will require you to make many decisions on your own. Your goal is to have a portfolio-ready case study. To make sure you are set up for success, it is recommended that you complete one of the more structured case studies first. Once your case study is complete, download the packet and reference the details of your work anytime to your online portfolio profile.

Scenario

You are a junior data analyst working for a business intelligence consultancy. You have been at your job for months, and your boss feels you are ready for more responsibility. He has asked you to lead a project for a brand new client — this will involve everything from defining the business task all the way through presenting your data-driven recommendations. You will choose the topic, ask the right questions, identify a fresh dataset and ensure its integrity, conduct analysis, create compelling data visualizations, and prepare a presentation.

Five questions will help you to guide your case study:

1. What type of company does your client represent, and what are they asking you to accomplish?
2. What are the key factors involved in the business task you are investigating?
3. What type of data will be appropriate for your analysis?
4. Where will you obtain that data?
5. What are the solutions you are trying to provide based on your business question?
6. Who is your stakeholder, audience, and what materials will help you present to them effectively?

Project outcomes

- Selecting and developing a case study: select and design your own case study about a topic that you are interested in. You can practice all phase of the Data Analysis Process: Ask, Prepare, Process, Analyze, Share, and Act (you will find more information in project step details)
- Practicing your case study pitch
- Having portfolio project ready and starting to creating your online portfolio after completing

Grading Information

For your project to be graded in subsequent steps, you will be required to upload the completed **Excel workbook**, **report** and **slide presentation** in google classrooms.

The details of grading information will be sent to you after your first draft submission. The main grading criteria will be:

- **Report (30%)** : The structure of your report and professional design, writing skill, Your report findings supported by strong statistical analysis and your analysis detail.
- **Dashboard (40%) :** dashboard appearance and design, calculation and interactiveness, flow, user friendly, correct visualization, excel formula and metrics used
- **Presentation (30%) :** how well storytelling and engaging, professional slide preparation and design, presentation criteria

You will not be judged on:

- Your English language, including spelling or grammatical mistakes in your report
- The content of any text or image(s) or where a link is hyperlinked to.

Guidelines for the Submission and deadline

For submission and deadline, you will need to check your google classroom assignment under the topic Excel Dashboard Project. The whole project will be broken down into small steps for your submission, so keep the track of your Excel Dashboard project assignment.

Software used in this project

This project is aimed at polish your excel skill, so you will use Microsoft Excel to do your analysis and build your dashboard. It is not limited to you if you want to use other software to be an add on to your project work such as BI tools.
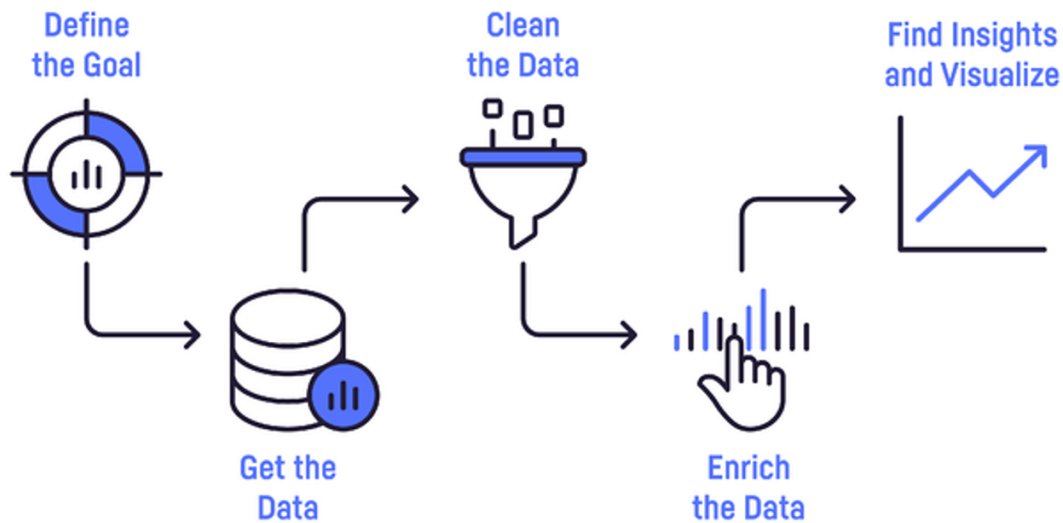
Dataset used in this project

You are free to choose the dataset based on your interest and industry focus. In your project step detail, you are given some sources to explore and find your interesting dataset. If you

DATA J

cannot decide which dataset or couldn't find the dataset, request our hardskill team to give you the dataset as we have already prepared for you.

Formatting Guidelines

- **Font Size**: In the report, 12 in the body text, except the topics and titles, which should be font size 14
- **Font Type**: Times New Roman or Arial
- **Spacing**: The project report should be 1.5 lines spacing
- **Highlighting**: Topics and subtopics should be bolded and NOT be underlined
- **Margins**: Margins of the report should be 1 inches on the top, bottom and right hand side. The left hand-side margin should be 1.25 inches to allow for binding.
- **Tables:** Larger tables may be typed in smaller font size (10-11) in order to maintain standard margins
- **Numbers and Percentages**: must not begin a sentence.
- **Tables and Figures** : When presenting the table or figure there must be a finding section and an analysis section. Avoid using 'table above, or table below.' Instead indicate as 'Table 4.1 shows that ….'
- **Pagination**: Bottom of page and centered.

DATA U
CREATING TOMORROW'S DATA LEADERS, TODAY

Flow Chart Indicating The Basic Elements For Project Work

Define the Goal

Get the Data

Clean the Data

Enrich the Data

Find Insights and Visualize

Project Checklists or detail:

1. Selecting and developing a case study

You will produce a report with the following deliverables:

1. A clear statement of the business task you have selected to investigate (analytic plan)
2. A description of all data sources used (dataset and data dictionary)
3. Documentation of any cleaning or manipulation of data
4. A summary of your analysis (your business solutions)
5. Supporting visualizations and key findings (dashboard)
6. Based on what you discover, a list of additional deliverables you think would be helpful to include for further exploration
7. Your top high-level insights based on your analysis (an executive summary)

Note: Completing this case study within a week is a good goal.

DATA J

Case study roadmap:

To begin, spend some time brainstorming ideas for your case study development. You will follow the roadmap below from Ask, Prepare, Process, Analyze, Share and Act. with each step of the roadmap, you are given the guiding questions, key tasks and deliverables. Use them as your checklist for your project case study development.

Step1 : Ask

Guiding questions

- What topic are you exploring?
- What is the problem you are trying to solve?
- What metrics will you use to measure your data to achieve your objective?
- Who are the stakeholders?
- Who is your audience?
- How can your insights help your client make decisions?

Key tasks:

1. Identify the industry of fucus, business task or case, business questions or Defining the topic of interest / business pain-points
2. Determine key stakeholders
3. Choose a dataset
4. Establish metrics
5. Creating analysis plan

Deliverable: A clear statement of the business task you have selected to investigate

Step2: Prepare

Find data that is appropriate for your analysis from any credible dataset. There are numerous public datasets that you can use. Once you have decided on a dataset, gather and store the data appropriately, clean the data, and make sure it is reliable and error-free. Document your process, as clients often ask to see both raw and cleaned data. Now, prepare your data for analysis using the following Case Study Roadmap as a guide:

Guiding questions

DATA J

- Where is your data located?
- How is the data organized?
- Are there issues with bias or credibility in this data?
- How are you addressing licensing, privacy, security, and accessibility?
- How did you verify the data's integrity?
- How does it help you answer your business questions?
- Are there any problems with the data?

Key Task

1. After defining the topic of interest, the focus shifts to the collection of fundamental data to elaborate the project, sourced from available databases. Another way to collect the necessary data revolves around establishing adequate connections to web APIs or collecting data directly from relevant websites with the potential for future analysis (web scraping).
2. Download data and store it appropriately.
3. Identify how it's organized.
4. Sort and filter the data.
5. Determine the credibility of the data.

Deliverable: A description of all data sources used and data dictionary

Note: there are some sources on the internet that provide an interesting dataset for you to choose. Here are the list of websites:

- IBM Data and AI Accelerators: https://community.ibm.com/accelerators/?context=analytics&type=Data
- Kaggle: https://www.kaggle.com/datasets
- Socrata: https://opendata.socrata.com
- Google Public Datasets: https://cloud.google.com/bigquery/public-data/
- Data.gov: https://www.data.gov
- Quandl: https://www.quandl.com/search

Follow these steps:

1. Download the dataset.

2. Create a folder on your desktop or Drive to house the files. Use appropriate file-naming 3. conventions. Create subfolders for .CSV or .XLS files.

4. Follow these instructions for either Excel (a) or Google Sheets (b):

- Launch Excel, open your file, and choose to Save As an Excel Workbook file. Put it in the subfolder you created for .XLS files.
- Open each .CSV file in Google Sheets and save it to the appropriate

5. subfolder. Open your spreadsheet and appropriately label the columns. 6. Proceed to the analyze step.

If you like, continue working with the data to better familiarize yourself and perhaps even identify new approaches to answering the business questions.

Step3: Process:

Then, process your data for analysis using the following Case Study Roadmap as a guide:

Guiding questions

- What tools are you choosing and why?
- Have you ensured your data's integrity?
- What steps have you taken to ensure that your data is clean?
- How can you verify that your data is clean and ready to analyze?
- Have you documented your cleaning process so you can review and share those results?

Key tasks

1. Check the data for errors.
2. The modification of data and columns will be necessary in order to confirm that no variables are missing.
3. Choose your tools.
4. Transform the data so you can work with it effectively.
5. Document the cleaning process.

Deliverable: Documentation of any cleaning or manipulation of data

Step 4: Analyze

Now that your data is stored appropriately and has been prepared for analysis, start putting it to work. Use the following Case Study Roadmap as a guide:

Guiding questions

- How should you organize your data to perform analysis on it?
- Has your data been properly formatted?
- What surprises did you discover in the data?
- What trends or relationships did you find in the data?
- How will these insights help answer your business questions?
- How will you organize your dashboard and choose the visualization?
- What does your dashboard look like?

Key tasks

1. Aggregate your data so it's useful and accessible.
2. Organize and format your data.
3. Perform calculations and KPI analysis
4. Create descriptive statistics and test several hypotheses – significant variables
5. Use regression formulas and algorithms to generate predictive models and foresee values and future patterns, in order to generalise occurrences and improve the efficiency of decisions
6. Document your calculations to keep track of your analysis steps.
7. Identify trends and relationships.
8. interpret the defined models and discover important business insights – finding generalisations to apply to future data – and respond to or address all the questions asked at the beginning of the project.
9. find patterns that can help companies in their decision-making processes: whether to avoid a certain detrimental outcome or repeat actions that have reproduced manifestly positive results in the past.
10. Your draft of the dashboard (can be your hand drawing or digital drawing)

Deliverable: A summary of your analysis and dashboard draft

Follow these steps for using spreadsheets or excel

Open your spreadsheet application, then complete the following steps:

Where relevant, make columns consistent and combine them into a single worksheet.

2. Clean and transform your data to prepare for analysis.
3. Conduct descriptive analysis.
4. Run a few calculations to get a better sense of the data layout.
5. Create a pivot table to quickly calculate and visualize the data.
6. Once you are working with several individual spreadsheets, merge them using the tool you have chosen to use to perform your final analysis, either a spreadsheet, Excel or BI tools. Export a summary file for further analysis.

Step 5: Share

Now that you have performed your analysis and gained some insights into your data, create visualizations to share your findings. You will be presenting to your client and other stakeholders, so visuals should be sophisticated and polished in order to effectively communicate your insights. Use the following Case Study Roadmap as a guide:

Guiding questions

- Were you able to answer the business question?
- What story does your data tell?
- How do your findings relate to your original question?
- Who is your audience? What is the best way to communicate with them?
- Can data visualization help you share your findings?
- Is your presentation accessible to your audience?

Key tasks

1. Determine the best way to share your findings.
2. Transfer the obtained data into data visualisation
3. Create effective data visualizations.
4. Present your findings.
5. Ensure your work is accessible.

Deliverable: Supporting visualizations and key findings (dashboard)

Follow these steps:

Take out a piece of paper and a pen and sketch some ideas for how you will visualize the data.

2. Once you choose a visual form, open your tool of choice to create your visualization. Use a presentation software,
3. Create and choose the right visualization for your data visualization, remembering that contrast should be used to draw your audience's attention to the most important insights. Use artistic principles including size, color, and shape.
4. Ensure clear meaning through the proper use of common elements, such as headlines, subtitles, and labels.
5. Refine your data visualization by applying deep attention to detail.

Step 6: Act

Now that you have finished creating your visualizations, act on your findings. Organize the deliverables you created, including your top high-level insights based on your analysis. Use the following Case Study Roadmap as a guide:

Guiding questions

- What is your final conclusion based on your analysis?
- How could your team and business apply your insights?
- What next steps would you or your stakeholders take based on your findings?
- Is there additional data you could use to expand on your findings?

Key tasks

1. Building your final analytic report
2. Practice presenting your case study to a friend or family member.

Deliverable

1. Your top high-level insights based on your analysis (data analytic report)
2. Based on what you discover, a list of additional deliverables you think would be helpful to include for further exploration

Other thing you should know beside the case study roadmap above

2. **Linear and logical spreadsheet flow-** Separate: Documentation, Dataset, Calculations, pivot calculation table, dashboard and Reporting. Create worksheets for each of these types of information and give the worksheets different colors based on the type of information in that worksheet. Each worksheet should only contain one kind of information (i.e data tabs should not contain calculations and vice versa).

3. **Protection & Hidden information-** Protect all sheets with formulas, make sure all formulas are locked so that they cannot be accidentally changed. Hide only whole worksheets, unhide all rows, columns, cells, ect. (this is optional, but we recommend you to do so to protect your work).

4. **Keep it Simple-** Keep your formulas short, you can split them into several cells. No formula should be longer than the formula bar. If you must create a formula longer than the formula bar, add a comment explaining the formula.

5. **Charts and Reporting-** If you do not already have a worksheet dedicated to the reporting part of the model, create one. Clearly represent the information in your model or analysis.

6. **Checking, Testing, and Validation-** Build in at least one automatic check, this can be as simple as a conditional format on a certain cell that goes red if the number is less than 0. You can also build more complex automated checks, but we understand this will not apply to all models, so we don't expect extensive checks.

7. **Finishing touches-**

   a. Remove all unused worksheets, named ranges, and styles.
   b. Remove any helper formulas or other temporary things you used along the way.
   c. Fix all standard Excel errors.
   d. Update all external links and database connections.
   e. Remove unused external links and database connections.
   f. Update all Pivot tables.
   g. Save your spreadsheet with cell A1 of your documentation/explanation worksheet selected.
   h. Give the spreadsheet a clear name with a version number.

If you meet all these requirements you instantly elevate the quality of your workbook..

2. Practicing your case study pitch

Presentation is also extremely important, as project results should be clearly outlined for the convenience of stakeholders (who, in the vast majority of instances, are without technical knowledge). The data scientist has to possess the "gift" of storytelling so that the entire process makes sense, meeting the necessary requirements to solve the company's problem. You will

learn more about presenting your project with Soft-skill sessions. With the information below is a summary related to presentation checklists.

- Prepare your slide presentation of your finding with the maximum presentation time of 20 minutes and be ready for 5 minutes Q and A.
- Present your finding in story telling way because your audiences will be non-technical people
- Make sure to keep your professionalism in your presentation. Bear in mind that you're not presenting an assignment findings but a business solution or business analysis result.