

PROJECT ASSIGNMENT #1

Brief

In this document you will find the description of the Group Project Assignment. Please, read it carefully.

You don't have to start working on the project right away. We will notify you via Discord when the time comes.

This project simulates a professional task that gives you a glimpse into the work of a real data analyst. You can add this project to your portfolio to get a significant advantage when you pursue a job in data analytics.

GENERAL INFORMATION

GOAL

The main objective of this project is creating a professional report in Excel, which is one of the most common tasks that data analysts handle on a daily basis.

You will be working in groups of approximately 6 people. To demonstrate that you are ready to tackle the responsibility of a real job, you will be: organizing your own workspace, hosting your own synchronous group meetings, discussing your work in private chats on Discord, etc. This format will prepare you for the job of a data analyst and will provide everyone in the group with an opportunity to contribute to the project and to prove themselves.

DEADLINE

You will have two weeks to complete the project. Your team will need to submit all the appropriate materials on the given deadline day.

PREREQUISITES

To participate in the project, you must:

1. Complete Module 2: submit assignments 2.3.9, 2.3.10, and 2.6.
2. Complete Module 3: submit assignments 3.3.7 and 3.3.6.

If you haven't completed these Modules yet, you won't be able to participate in this project, because you require both SQL and Excel knowledge to do it. Please, register for this project after you complete the required modules.

GENERAL INFORMATION

BACKGROUND

You are a data analyst in the company Yellevate, which specializes in providing marketing services to other companies. Yellevate helps mid-sized companies to launch their marketing operations, which includes things like email marketing, website development, content creation, and others.

Yellevate uses data analytics to keep track of how well it helps companies improve their marketing operations. For instance, if a company approaches Yellevate to help improve its online advertising operations, Yellevate tracks this by looking at email impressions, click-through rates, website traffic, and other key performance indicators.

TASK DESCRIPTION

You and your team in the data analytics department were approached by the company's top management to solve a serious problem. For the past few years, Yellevate has been struggling with client disputes. Yellevate defines disputes as clients expressing dissatisfaction with the company's services and refusing to pay for them.

This has been a huge financial burden for the company: statistically, nearly 20% of the disputes raised against Yellevate resulted in a payment opt-out. This has led to an approximate 5% annual loss of revenue (in USD).

Management has now approached you and your team to help resolve the issue. Using data the company collected about these disputes, you will have to identify their causes and come up with actionable strategies to solve them. The analysis you present will be used by the company's top executives, so you need to absolutely make sure that you provide an exceptional report: after all, there are millions of dollars on the line.

DATA ANALYSIS GOALS

Executives at the company decided that you need to obtain the following information to identify the circumstances around the dispute problem:

1. The processing time in which invoices are settled (average # of days rounded to a whole number).
2. The processing time for the company to settle disputes (average # of days rounded to a whole number).
3. Percentage of disputes received by the company that were lost (within two decimal places).
4. Percentage of revenue lost from disputes (within two decimal places).
5. The country where the company reached the highest losses from lost disputes (in USD).

ACTION PLAN

Step #1. Identify the problems facing the business and come up with objectives that can solve them with data analysis.

Step #2. Set data analysis goals according to these business objectives.

Step #3. Conduct your data analysis by loading a .csv file in SQL to perform the necessary data cleaning.

Step #4. Analyze data using Excel functions.

Step #5. Visualize your processed data in a manner that your visuals will help answer any and all the questions.

Step #6. Generate insights from your analysis to provide recommendations on probable strategies to deal with disputes effectively.

ACCESSING THE DATA

All the data you will be processing for this project can be accessed in [Yellevate Invoices.csv](#)* and its [data dictionary](#)*.

* *This dataset was created for this project specifically.*

IMPORTANT ASSUMPTIONS**

Keep this information in mind as you develop your insights and recommendations:

1. All the services Yellevate was hired to do were fully completed.
2. The quality of the services provided is not the main driving reason for these invoice disputes.
3. Yellevate management believes most disputes are the result of contract technicalities or clients thinking they can get away with not paying for the services.

refocus

*** An assumption is something that you accept as true without question or proof.*

ADDITIONAL INFORMATION

When importing data into SQL, you can use this list of data types expected for each variable as a guide to help:

country: varchar
customer_id: varchar
invoice_number: numeric
invoice_date: date
due_date: date
invoice_amount: numeric
disputed: numeric
dispute_lost: numeric
settled_date: date
days_settled: integer
days_late: integer

DOCUMENTS TO HAND IN

#1. REPORT ON THE ISSUE

Student #1, Student #2, Student #3, Student #4, Student #5, Student #6
Date of submission: XX.XX.XXXX

Project Title

Problem

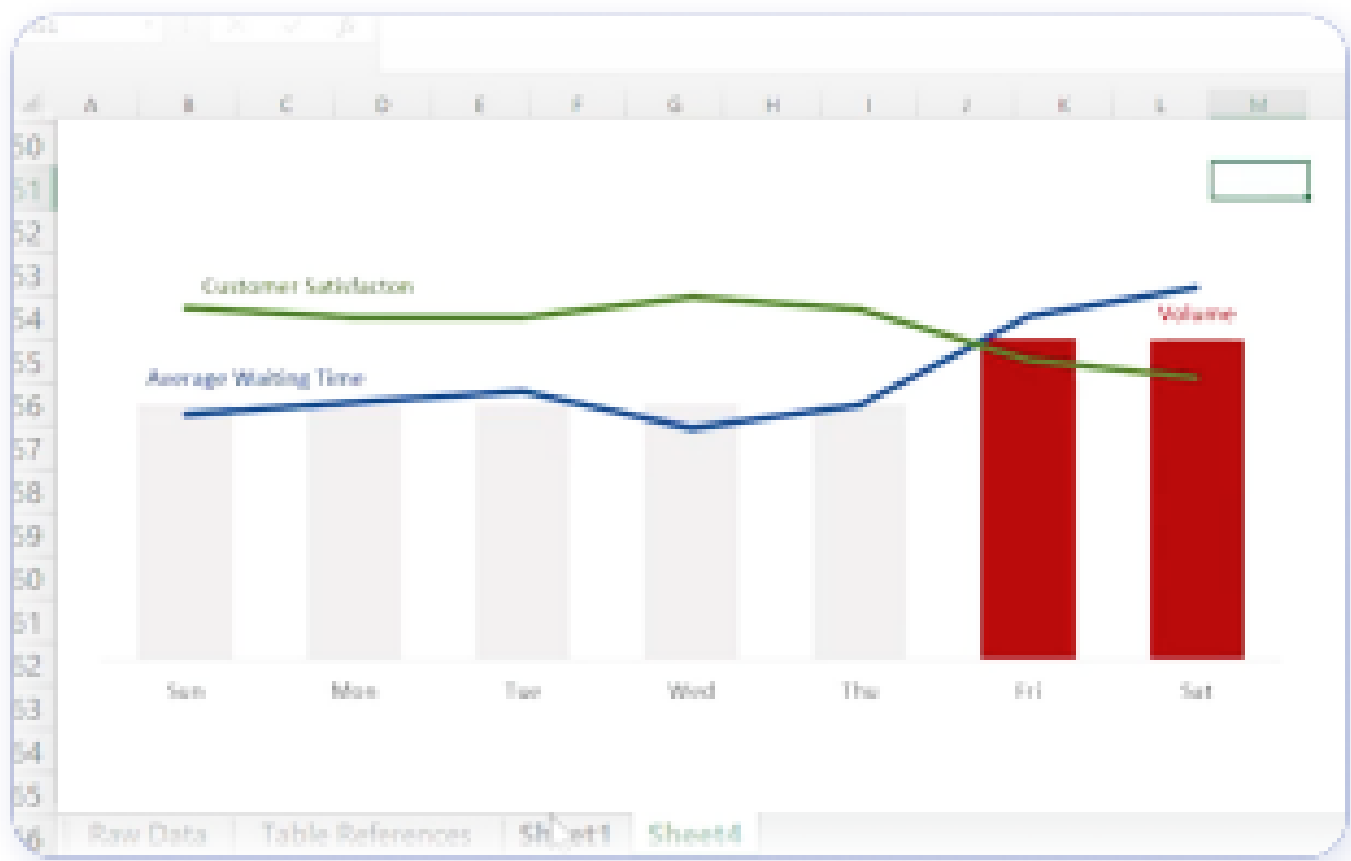
The problem was...

Methodology

That's how we investigated the issue...

Findings

Here are insights that we gathered through data analysis...



Recommendations

That's how we recommend solving the problem...

Pic. 1 An example of how your report can look

File Format: Word document converted into PDF.

First Page should include:

- ★ project title
- ★ student names
- ★ date of submission

refocus

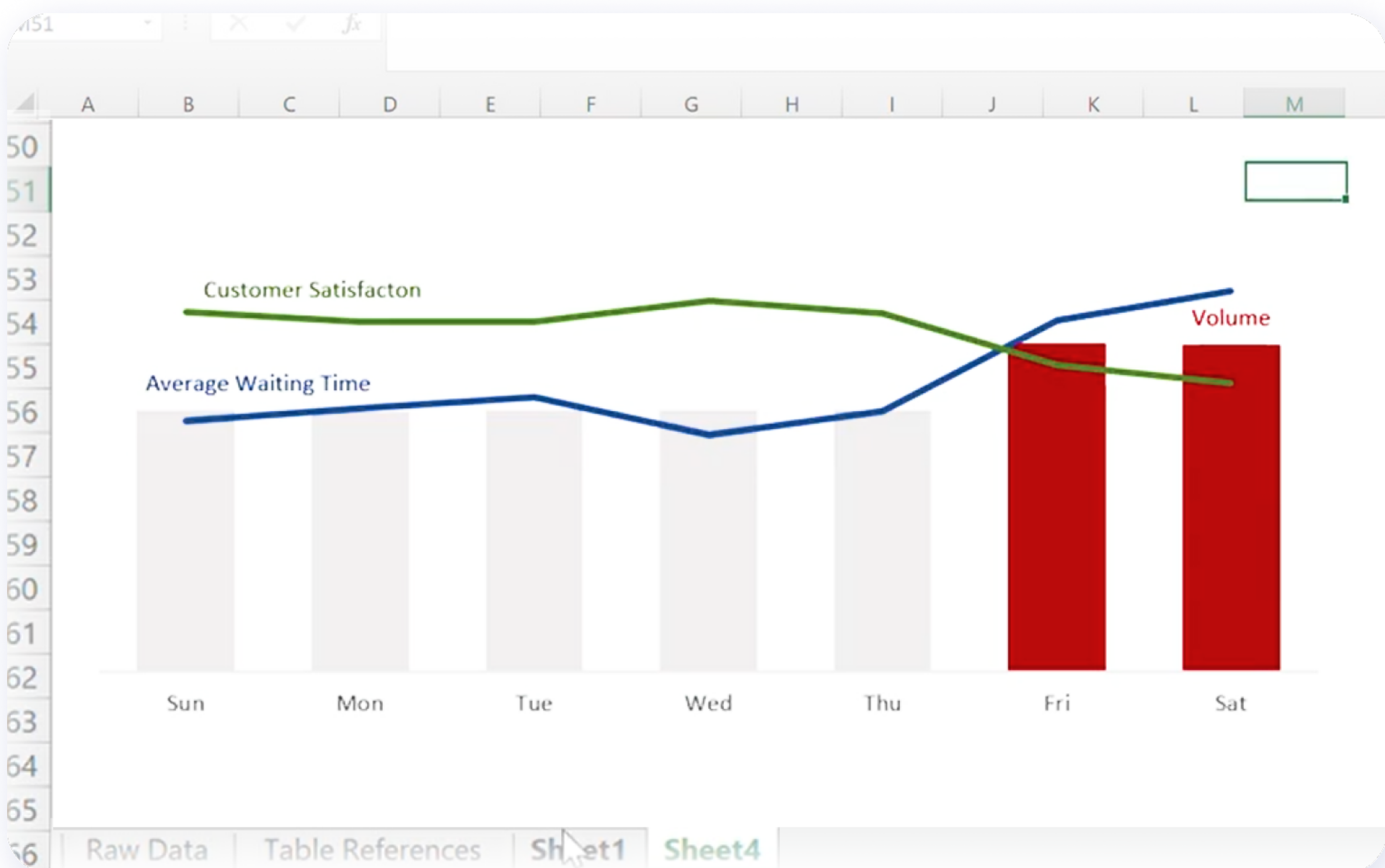
Your report should provide:

- ★ written insights
- ★ the overall situation with client disputes
- ★ recommendations to fix the situation
- ★ visuals and data tables from your Excel file that support all your comments

#2. PROCESSED DATA FILE IN EXCEL AND SQL

This file should include: all the processing you have performed on your raw dataset to get your results.

Make sure to perform all your visualization here as well. This is a professional report you would provide to a company's top management, so make sure it looks neat and well-organized.



Pic. 2 An example of how visualization in your data file can look

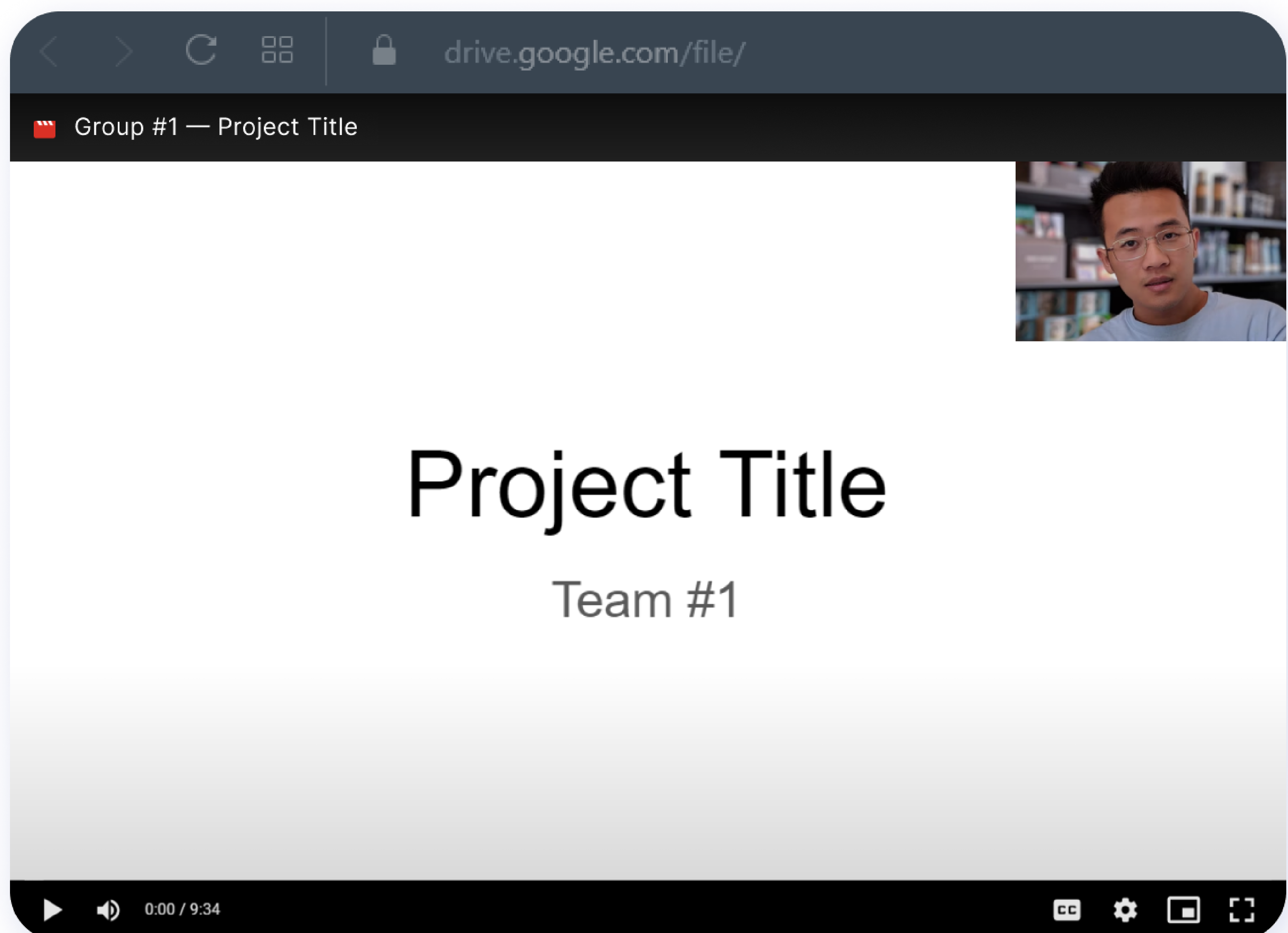
#3. RECORDED VIDEO PRESENTATION

Imagine your group is being called in to present your findings to the top management. Structure and present your findings using a PowerPoint presentation.

Your presentation should include:

- ★ explanation of all the work you did
- ★ the results you found
- ★ the actionable strategies you suggest to effectively deal with disputes

Make a video of your group presentation using the recording software of your choice. For reference, you can use Zoom to record your video. Instructions for Zoom video recording [can be found here](#). Submit the video file along with the rest of your materials.



Pic. 3 An example of how your presentation can look

#4. LINK TO YOUR GROUP WORKSPACE

This is a document where you will need to track your group progress: who is responsible for each task, status and deadlines. This is essentially important while working in a group. Make sure it reflects how your team worked.

Student #1, Student #2, Student #3, Student #4, Student #5, Student #6
Date of submission: XX.XX.XXXX

Group Workspace

Task

Copy here the task points you think are the most important.

Plan

Task	Who is responsible	Deadline	Status
Conduct a meeting and divide the tasks	Joshua — Group Lead	XX.XX.XXXX	Completed ▾
Clean the data	Nathalie	XX.XX.XXXX	In Progress ▾
Calculate the processing time in which invoices are settled	Angelo	XX.XX.XXXX	Not Started ▾

Pic. 4 An example of how group workspace can be organized