TABLE OF CONTENTS

YELLEVATE

PROJECT ASSIGNMENT 1

PROBLEM

Yellevate, a company providing marketing services, has been struggling with client disputes. These disputes were defined to the analysts as clients expressing dissatisfaction with the company's services and refusing to pay for them.

The company's problem is the 5% annual loss of revenue (in USD) due to losing 20% of the disputes raised by clients.

The analysts are tasked to identify the causes of these disputes and come up with recommendations and strategies to resolve them.

METHODOLOGY

The quantitative data the analysts used was collected by and provided by the client, Yellevate. The data was provided in a *csv* file and a Data Dictionary was also provided in *xlsx* format. The data was prepared and analyzed with the following steps:

1. The analysts used **PostgreSQL** as the database management system for the data and pgAdmin 4 was the user interface used. Data was cleaned and computations for metrics were done using this interface.
2. After data cleaning, the table was exported to **MS Excel** for further analysis using Pivot Tables and Visualizations.

3. Insights into the problem and recommended solutions were provided using the Visualizations and the **Underlying Assumptions** provided by the client as follows:

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

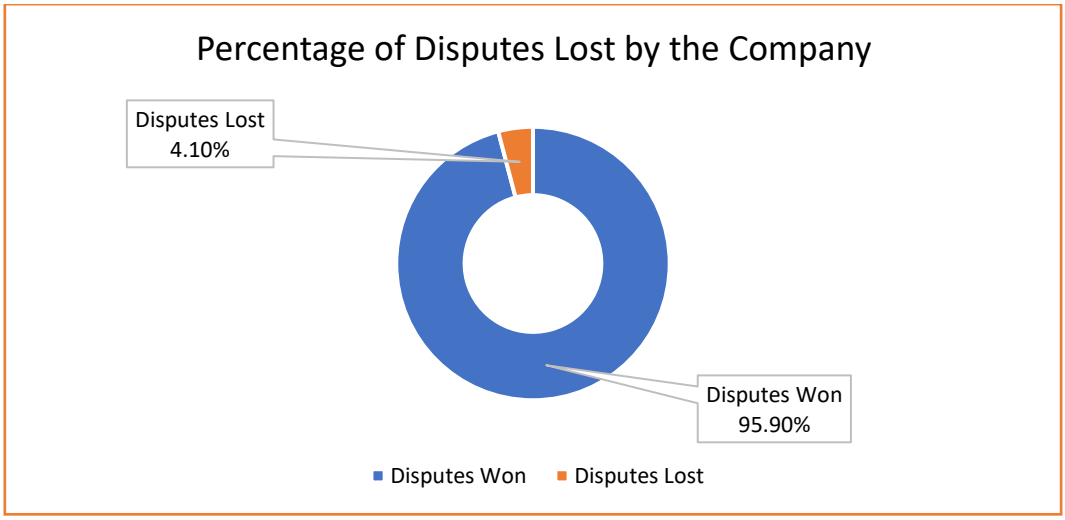
4. The following data analysis goals, calculated using Pivot Tables and SQL queries were also used to gain insights into the problem:

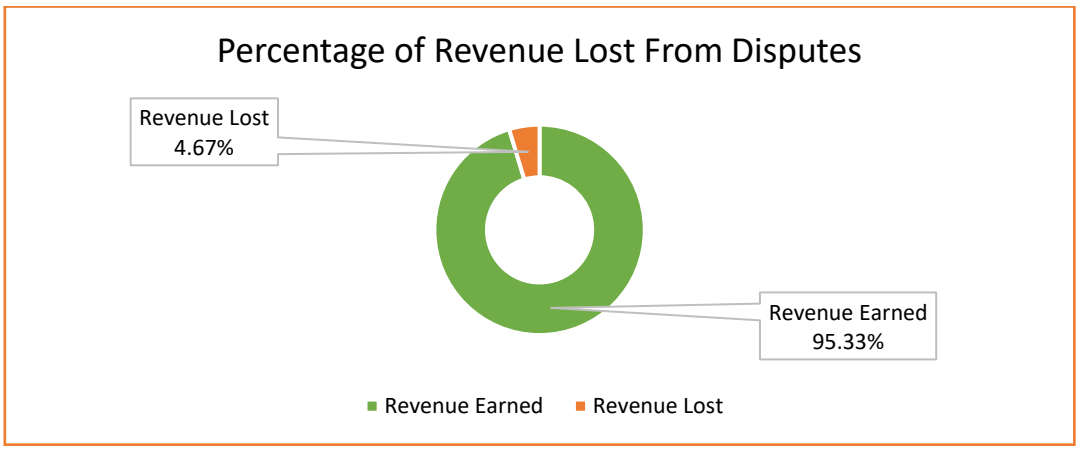
    a. The processing time in which invoices are settled (average number of days rounded to whole number).
    b. The processing time for the company to settle disputes (average number of days rounded to whole number).
    c. Percentage of disputes received by the company that were lost (within two decimal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

FINDINGS

I

Based on queries executed with PostgreSQL using pgadmin, the percentage of disputes lost by the company, rounded to two decimal places, is **4.10%**. The percentage of revenue lost by the company, rounded to two decimal places, is **4.67%**.

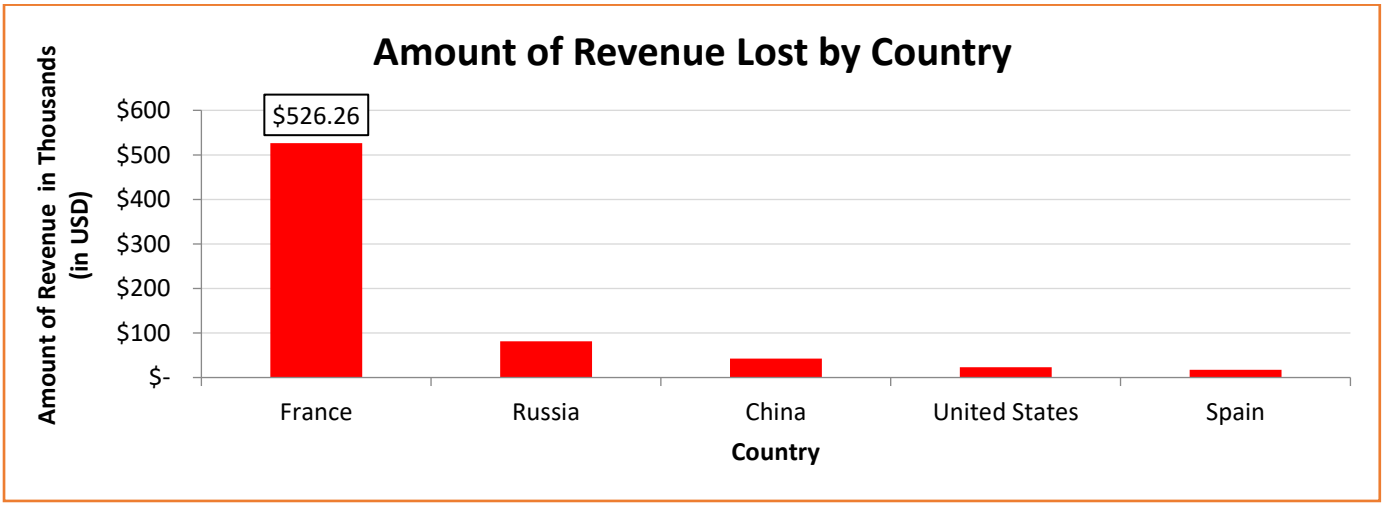


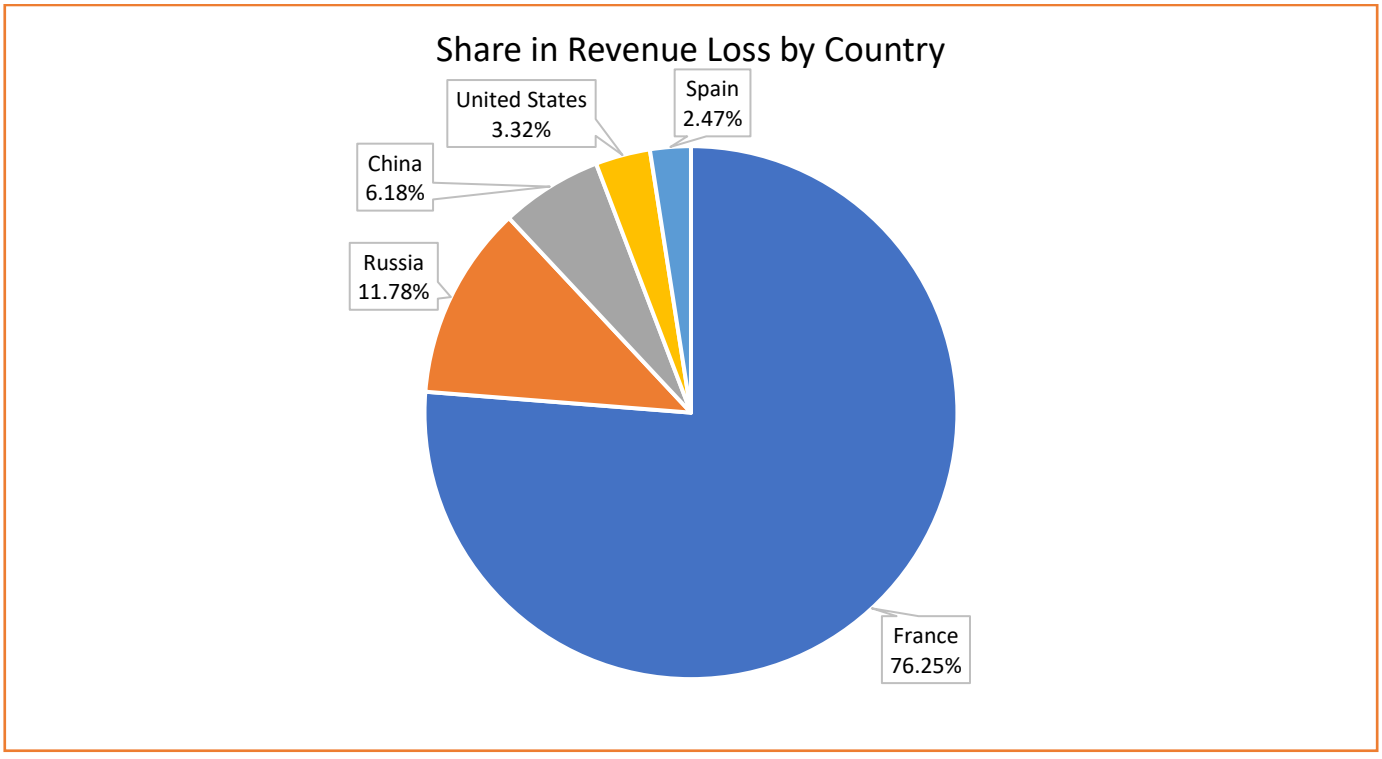


At a glance, 4% to 5% a loss does not seem too bad, but when these losses add up every year, it is no longer an insignificant amount.

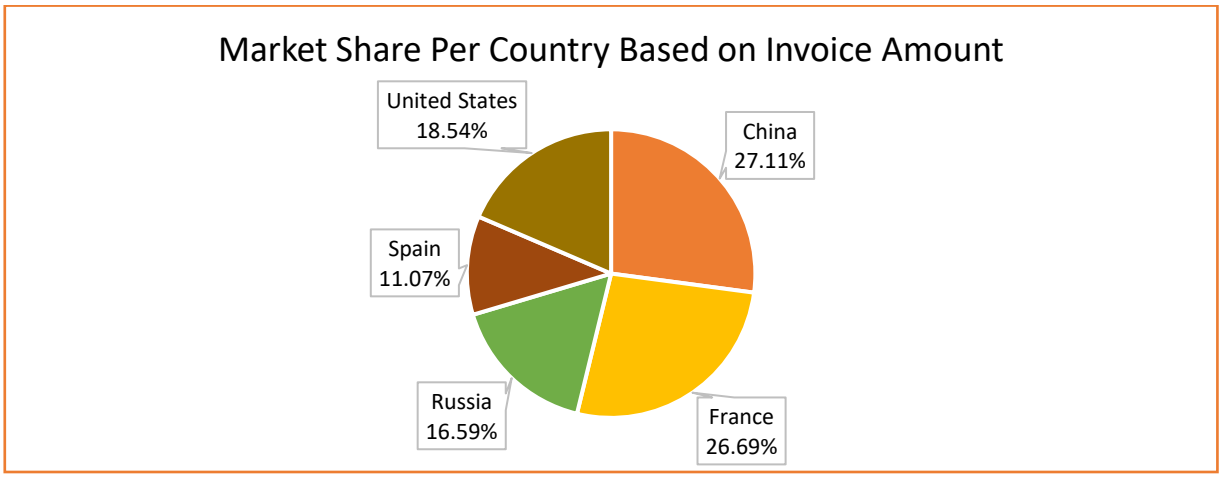
II

Objectively, we are trying to find where these lost disputes are coming from. Based on the charts below, we could see that our market in France has the highest amount of lost disputes with around $526,000 of losses which makes up 76.25% of the total revenue lost. This alarmingly high amount and percentage share of losses must be addressed and remedied.





Looking at the chart below, we can see that France has the second largest market share of our services based on total invoices. Thus, the option to disregard the market of France is not viable.

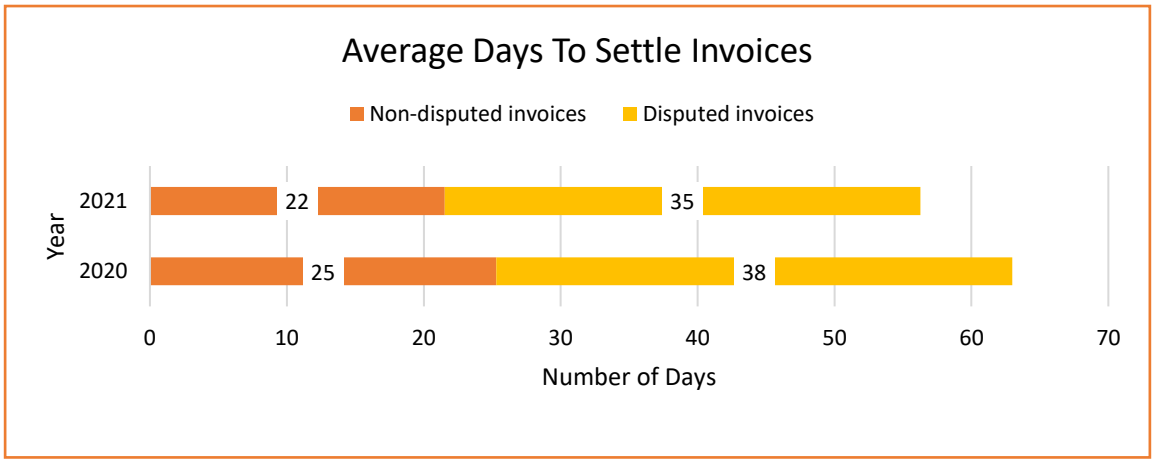


III

When it comes to settlement of invoices, the average number of days for invoices to be settled is 26 days. For disputes, the average number of days to settle disputes is 36 days. Both these averages decreased from 2020 to 2021 which does seem to be positive, as this would mean less frustration for clients. However, taking a month to settle disputes seems like too long a wait.

	average_processing_time numeric
1	26

	average_disputed_processing_time numeric
1	36

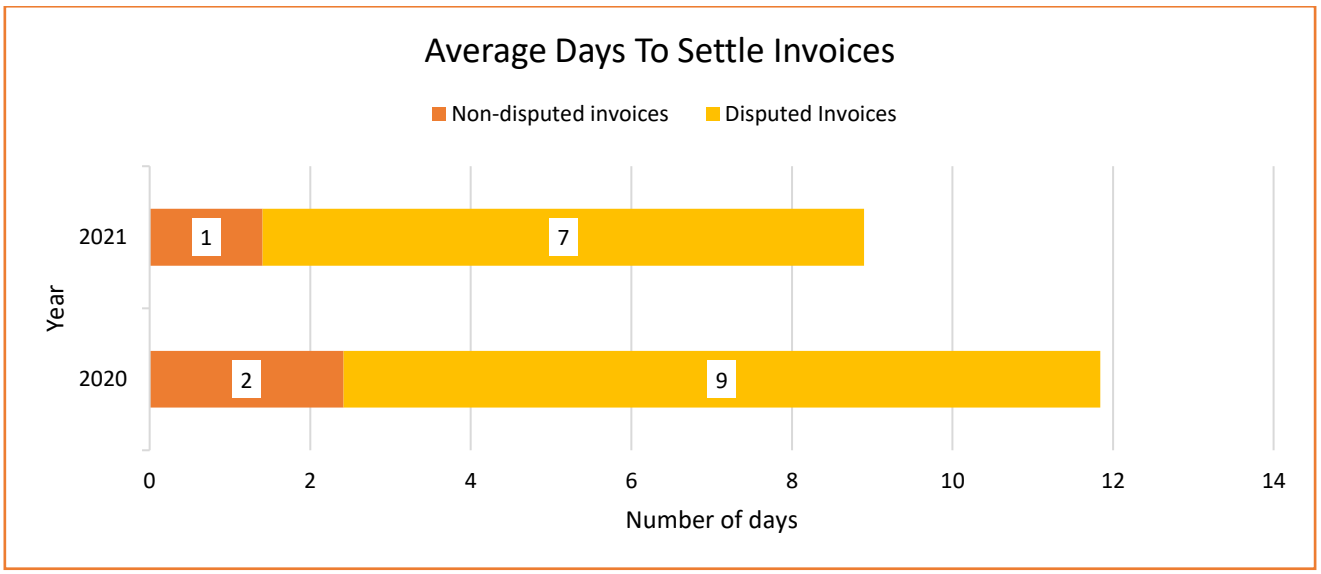


IV

As for late settlement of invoices, the average for non-disputed invoices is 2 days, while disputed invoices can be settled late up to 8 days. However, credit must be given to the company as it has reduced this average from 2020 to 2021.

Ideally, we are aiming for zero days late settlement of invoices and disputes.

Average number of days settled late	Invoices		
Total	Non-disputed invoices	Disputed Invoices	Grand Total
Grand Total	2	8	3



RECOMMENDATIONS

Taking the processed data and underlying assumptions into consideration, the analysts have come up with several recommendations relating to:

I. contract technicalities
II. payment terms
III. pricing
IV. incentives/penalties for early/late payment
V. more effective dispute resolution process
VI. better customer service

I

When it comes to **contract technicalities**, the analysts recommend that the company set up a legal department with multiple teams specializing in this matter. This legal department must properly review and revise the contract terms and conditions, clarify expectations, and responsibilities to reduce disputes.

The company can set up teams based on:

a. the type of services the company offers, or;
b. per country to which the company offers its services

They can have teams that specialize in international law and contracts, and/or some other teams that are well versed in contracts per country. If having multiple teams is too costly to the company, they can hire external lawyers or law firms to help them with the legal technicalities. The company

would simply have to choose between paying employees or paying retainer's fees on a periodic basis.

Seeing as France has the largest share in lost disputes and the second largest market share, we recommend focusing on this country first and thoroughly check the contracts for this location.

II

Clients might think they can get away from paying because of **unclear payment terms or invoices**. The company could discuss with the client the terms of payment during the contract signing process. Important details must be disclosed in the invoices, such as:

- invoice price
- if payments are in installment, the payment dates, payment amounts, interest rates, if any
- breakdown of costs, if any
- termination fee

III

The analysts also recommend studying other competitors in the country and industry to **price** the company's services more competitively and more affordably to its clients.

IV

Having **incentives or credits** for prompt payment and **penalties** for late payment might also be beneficial. This could help reduce the average number of days for a client to settle their payment. Some of these incentives or penalties may be:

- discount or bonus service for clients who consistently pay on time
- discount for lump sum payment

- discount or bonus service for clients who continuously avail of services through Yellevate (for example, every 3 or 5 services, they get a discount on their next service fee/invoice)
- a penalty or surcharge for late payment of invoices (if on installment basis)

V

From the data presented earlier in the insights, the average number of days for clients to get their money back is 36 days. A month to settle disputes can be too long for a client, and the company should aim to reduce this. Not only will accomplishing this encourage loyal clients, but it will also attract more customers. The fact that there are also late settlement of invoices, especially disputed ones, is a fact that needs to be addressed.

The analysts recommend implementing a **clear dispute resolution process**s that will clearly outline the steps when a dispute arises. This will help provide clear timelines to help reduce the average time in settling disputes. This may include having a handbook or conducting seminars for employees that are handling the dispute process.

VI

Varied **customer service channels** must also be made available, if possible. Clients should be able to contact the company through email, phone call, text, or virtual agents. Customer service should be easy to reach and there should be little to no waiting time.

Alternatively, to ensure that Customer Service teams are doing their job effectively and efficiently, practical performance measures and performance monitoring should be done on a periodic basis. However, the company should note that the measure of performance is **practical/realistic** and not **idealistic.**

While Yellevate is aiming to retain old clients and attract new clients, it should remember to treat its employees humanely and compensate them equitably.

-END-

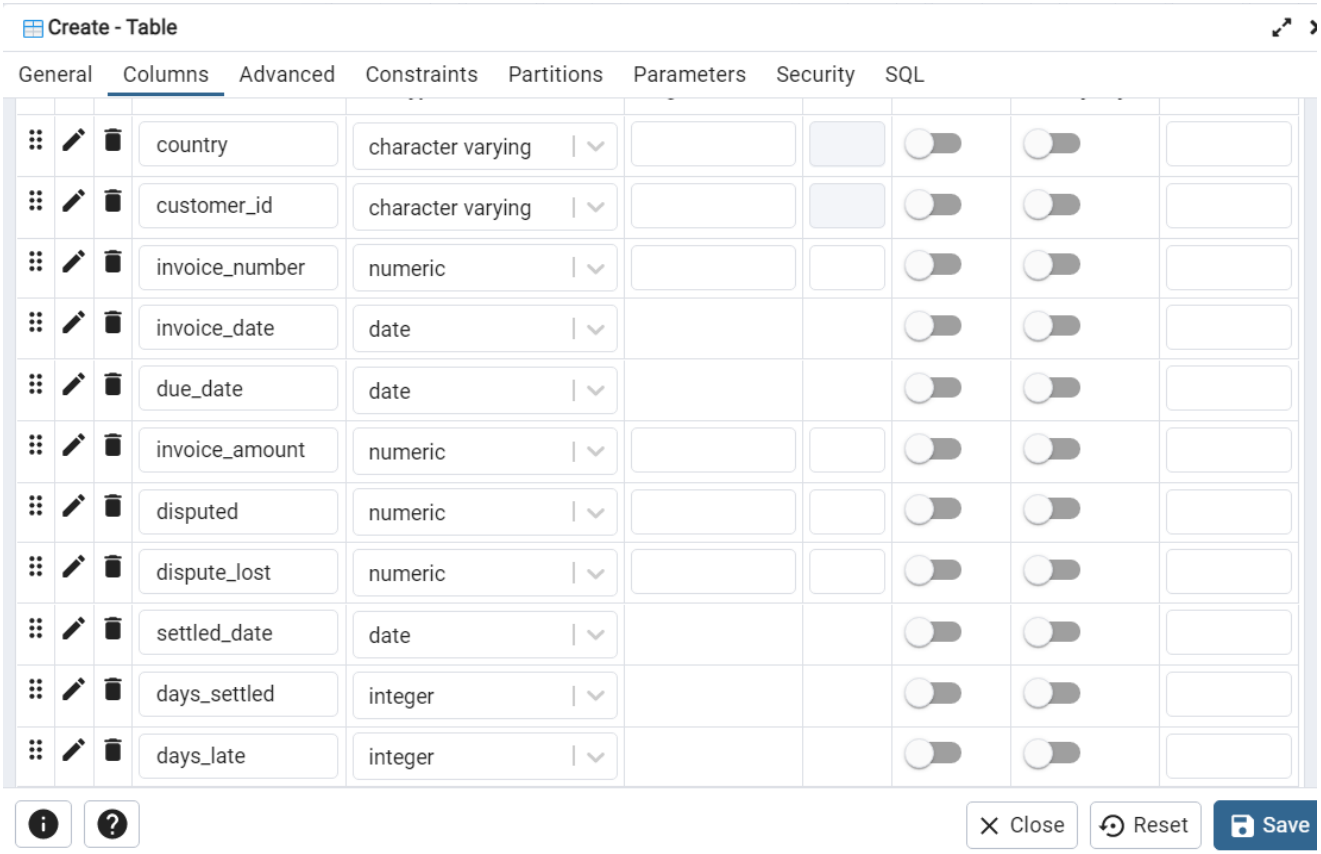
Please see next page for appendix of SQL Queries

APPENDIX

SQL QUERIES

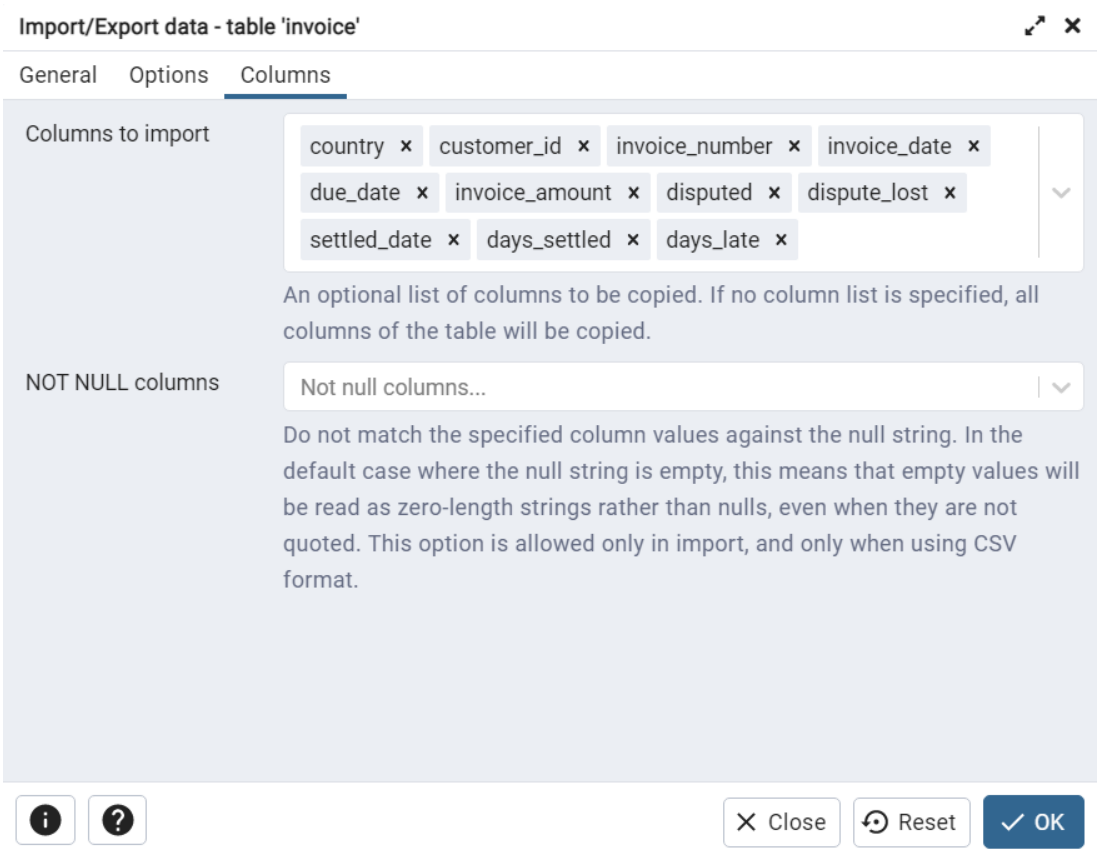
RESTORING AND IMPORTING DATASET

1. Created a database named “gp1_yellevate”.
2. Right-clicked Tables then Create < Table.
3. Inputted the table name in the General tab. For the project, we used “invoices”.
4. Navigated through the Columns tab and inputted manually the column headers and data type of fields in the CSV file. It should look like the image below.



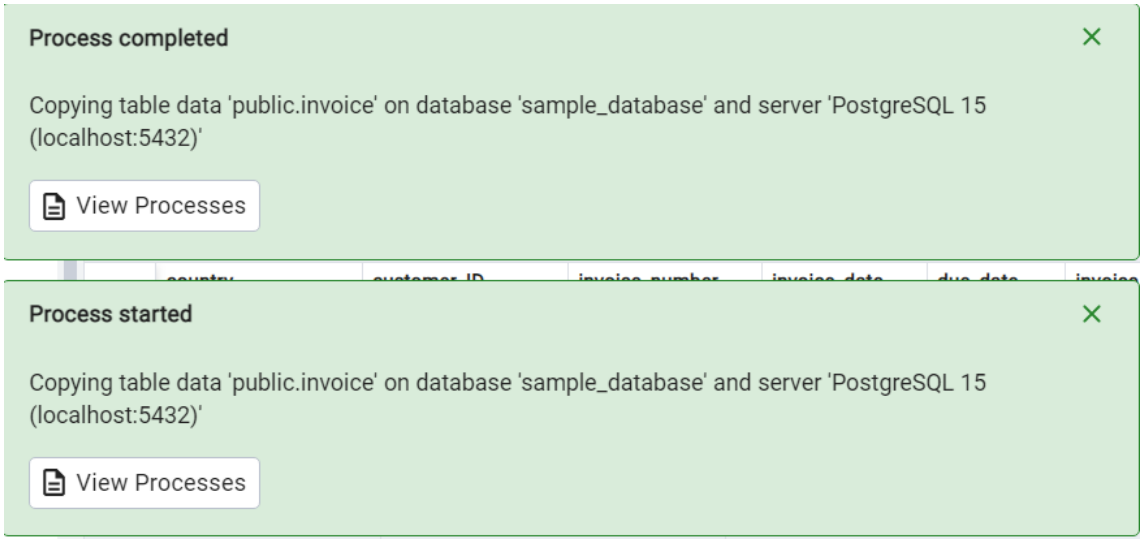
5. Click Save.
6. Right-clicked the created table name then proceeded to the Import/Export Data option.
7. In the General tab, choose Filename (CSV file of dataset) then remove OID in the Options tab.
8. In the Columns tab, the headers of the fields should be available in the Column to import section.

   It should look like the image below.



9. Pressed OK and it should return a Process Started and Process Completed notification.

   It should look like the image below.



CLEANING THE DATASET

In cleaning the dataset, the following were executed to determine outliers, duplicates, and nonsensical values:

B.1 Checked erroneous data

Checked whether there is a “1” value in dispute_lost then “0” in disputed in the same row. It should return with NO records.

SQL QUERY :

```
-- Checking if there are erroneous data where there is dispute_lost but is
not disputed
-- Should have NO data
SELECT disputed, dispute_lost
FROM invoices
WHERE disputed = 0
AND dispute_lost = 1;
```

RETURNED OUTPUT :

	disputed numeric	dispute_lost numeric

B.2 Checked possible misspelled or duplicated data in the country field.

SQL QUERY :

```
-- Checking wrong spelling in country

SELECT country

FROM invoices

GROUP BY country;
```

RETURNED OUTPUT :

	country character varying
1	Spain
2	Russia
3	China
4	United States
5	France

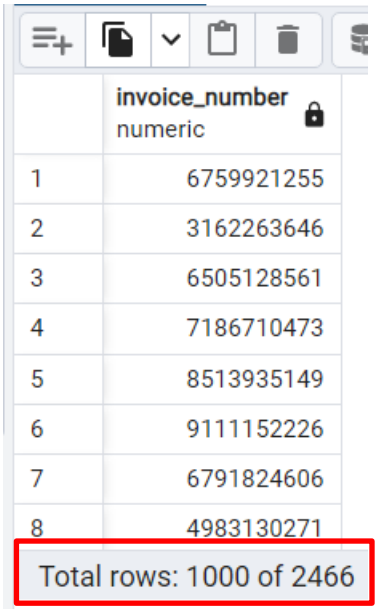
B.3 Checked possible duplicates in the invoice_number field.

The field invoice_number is deemed to be a unique identifier and cannot be duplicated. It should return the same number of rows as the total records if it is clean. Otherwise, remove duplicates

SQL QUERY :

```
-- Checking duplicates in invoice_number
-- Should be the same total number of rows as SELECT * FROM invoices;
-- Invoice number is a unique identifer
SELECT invoice_number
FROM invoices
GROUP BY invoice_number;
-- Returning all records
SELECT *
FROM invoices;
```

RETURNED OUTPUT :



	invoice_number numeric
1	6759921255
2	3162263646
3	6505128561
4	7186710473
5	8513935149
6	9111152226
7	6791824606
8	4983130271

Total rows: 1000 of 2466

	country character varying
1	China
2	France
3	China
4	France
5	Russia
6	Russia
7	Spain

Total rows: 1000 of 2466

CONCLUSION

After executing all procedures above, apparently, there were no errors found or missing data that needed to be replaced or filtered out. Hence, data is already clean.

SQL QUERIES FOR ANALYZING THE DATASET

The Processing Time In Which Invoices Are Settled

SQL QUERY :

```
-- Processing time in which invoices are settled. Either with or without disputes
SELECT ROUND(AVG(days_to_settle),0) AS average_processing_time
FROM invoices
```

RETURN OUTPUT :

	average_processing_time numeric
1	26

The Processing Time For The Company To Settle Disputes

SQL QUERY :

```
-- Processing time for the company to settle disputes
SELECT ROUND(AVG(days_to_settle),0) AS average_disputed_processing_time
FROM invoices
WHERE disputed = 1;
```

RETURN OUTPUT :

	average_disputed_processing_time numeric
1	36

Percentage Of Disputes Received By The Company That Were Lost

SQL QUERY :

```
-- Percentage of disputes received by the company that were lost
SELECT (
	SELECT COUNT(*)
	FROM invoices
	WHERE disputed = 1
	AND dispute_lost = 1
	) AS total_invoices_disputed_lost,
	COUNT(disputed) AS total_invoices,
	ROUND(
		ROUND(
			(SELECT ROUND(COUNT(*),4)
			FROM invoices
			WHERE disputed = 1
			AND dispute_lost = 1)
			/
			(SELECT ROUND(COUNT(*),2)
			FROM invoices)
		,4)
* 100,2) AS percentage_of_disputed_lost_to_disputed
FROM invoices;
```

RETURN OUTPUT :

	total_invoices_disputed_lost bigint	total_invoices bigint	percentage_of_disputed_lost_to_disputed numeric
1	101	2466	4.10

Percentage Of Revenue Lost From Disputes

SQL QUERY :

```
-- Percentage of revenue lost from disputes

-- Based on task description should be around 5%

SELECT (

	SELECT SUM(invoice_amount_usd)

	FROM invoices

	WHERE disputed = 1

	AND dispute_lost = 1) AS total_amount_disputed_usd,


	SUM(invoice_amount_usd) AS total_amount_usd,


	ROUND(

		ROUND(

			(SELECT ROUND(SUM(invoice_amount_usd),4)

			FROM invoices

			WHERE disputed = 1

			AND dispute_lost = 1)

			/
```

```
                    (SELECT ROUND(SUM(invoice_amount_usd),2)
                    FROM invoices)
                ,4)
        * 100,2) AS percentage_of_disputed_lost
FROM invoices;
```

RETURN OUTPUT :

	total_amount_disputed_usd numeric	total_amount_usd numeric	percentage_of_disputed_lost numeric
1	690167	14770318	4.67

The Country Where The Company Reached The Highest Losses From Lost Disputes

SQL QUERY :

```
-- Country where the company reached the highest losses from lost disputes (in USD)
SELECT country, SUM(invoice_amount_usd) AS disputed_amount_lost
FROM invoices
WHERE disputed = 1 AND dispute_lost = 1
GROUP BY country
ORDER BY SUM(invoice_amount_usd) DESC;
```

RETURN OUTPUT :

	country character varying	disputed_amount_lost numeric
1	France	526264
2	Russia	81291
3	China	42630
4	United States	22936
5	Spain	17046