*Genevieve Gonzales, Eddie Galarce Jr,Kristian Glorioso, Neil Angelo Martinez, Onofre Vidad*

*Date of Submission: December 5,2022*

**PROJECT ASSIGNMENT #1**

**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. The team will be following Google’s five phases of data analysis: ask, prepare, process, analyze and share.

Statistically, approximately 20% of the Yellevate conflicts resulted in a payment opt-out. This has resulted in a 5% annual revenue loss (in USD). Management has now asked the data analyst team to assist in resolving the problem. The team must identify the reasons of these issues using data gathered by the company and devise practical methods to resolve them.

Executives at the company decided that the following information should be obtained 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).

**METHODOLOGY**

ROCCC Analysis:

Reliable – The data has been made available by Refocus specifically for this project

Original - Subject to the terms and conditions of the Data License Agreement by Yellevate.

Comprehensive – Data dictionary has been provided through this [link](https://docs.google.com/spreadsheets/d/1bm4UvkcBfIS-blftRSafsg1TWTyVZ3SD/edit?rtpof=true&sd=true#gid=503846598).

Current – The latest data is 2022.

Cited - All the data needed for this project can be accessed through this [link](https://drive.google.com/file/d/1DXOaFAFF1Rqic1PWXvABxgCS4o1gYgxG/view).

**CLEAN DATA**

According to the data dictionary, the *disputed* field has 0 or 1 values. "1" signifies the customer disputed the invoice; "0" means they did not. When this table is provided to clients, they must understand what the numbers 0 and 1 signify. The team has added a new column that will display Disputed or Not Disputed.

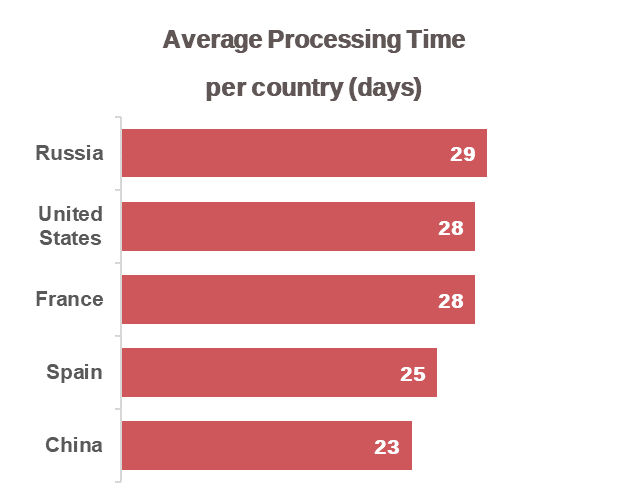
Same goes with the *dispute\_lost* column, "1" means Yellevate lost the dispute and the dispute was resolved in favor of the customer, and the customer does not have to pay the invoice; "0" means the customer did not win the dispute, and they are legally required to pay the full invoice amount, it either means that Yellevate won the dispute, or there was no dispute in the first place.

Lastly, a new column named *settlement* has been added to distinguish disputed that have been settled on time or settled late.

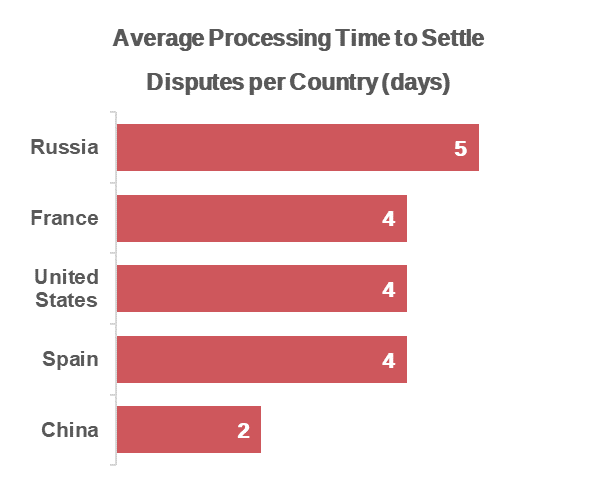
The country column is the one that can be checked for inconsistencies. We can see all of the countries listed by using the DISTINCT function. There are just five distinct countries after running the code.

**FINDINGS**

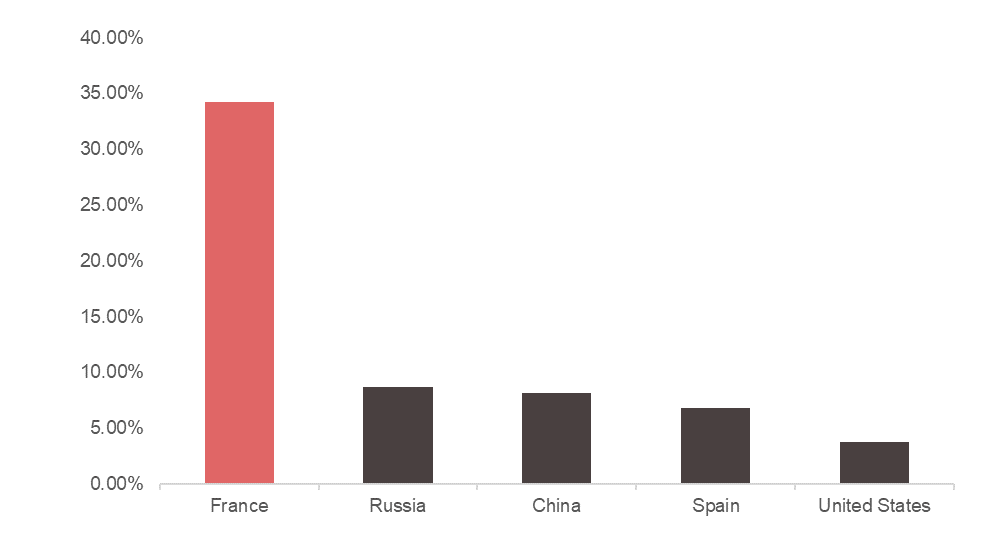
Russia has the longest average processing time of 29 days, which is 22.81% longer than China, which has the shortest data of 23 days. Furthermore, the average processing time for the entire dataset is 26 days.



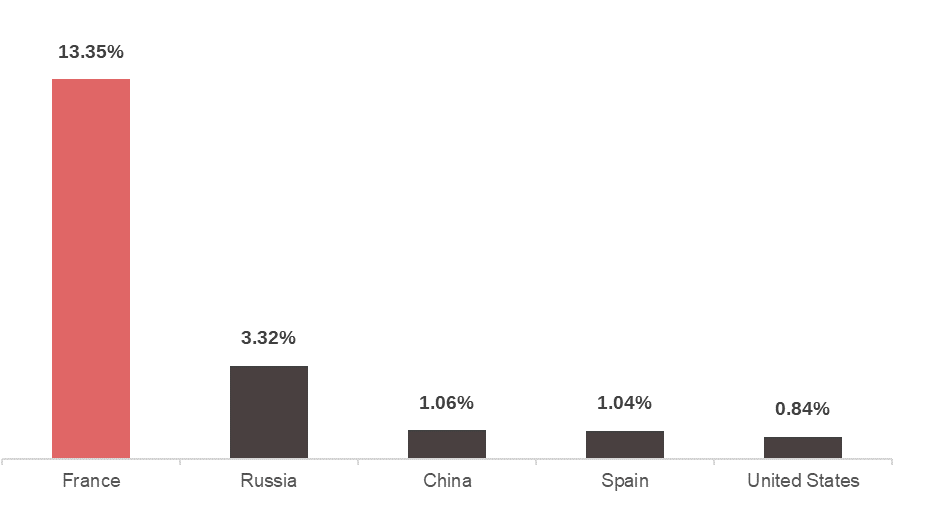
Russia has the highest average processing time of 5 days, which is 2.5 times faster than China, which has an average processing time of 2 days. Furthermore, the average processing time for the five countries is three days.



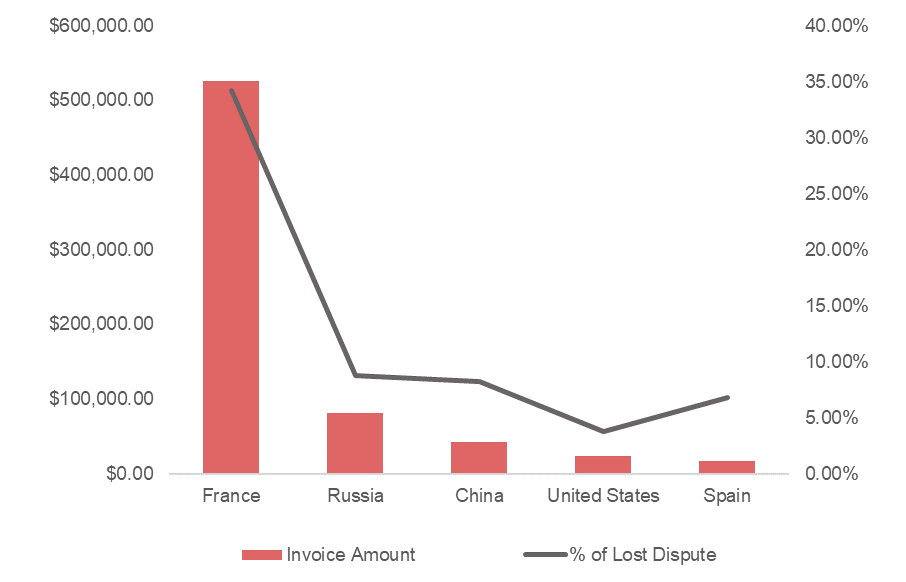
France has the highest percentage of lost disputes of 34.23% which is 9 times higher than United States which only has 3.75% lost disputes.



France has the highest percentage of revenue lost from disputes of 13.35% which is 15 times higher than United States which only has 0.84% lost disputes.



Because France has the highest rate of lost disputes, the country with the biggest losses from lost disputes should be France.



Customers file a dispute with Yellevate because they expected a pandemic exception in their contract but were not given one. As a result, they decided not to pay their invoices or file a dispute. They believe that because of natural disasters, every company should make an exception for their customers.

Consider that the invoice dates occurred during the pandemic, which meant that employees were adjusting from their workforce. Furthermore, the lack of documentation from both parties, and the lack of communication caused the dispute to arise.

Aside from the invoices being settled late, there is no visibility into why they were late and why the invoices being disputed. It is suggested that we include more information in the dataset, such as the cause for the delay and why it is disputed..

With it, we can get a better understanding of the issue at hand and take the required steps to resolve it. For the time being, since the only ground for a dispute is "Late Settlement," we can ask the client how the settlement is progressing.

**Recommendations**

Yellevate and their customers must have an open communication with each other especially if one party wanted to change the contract.

Have a discussion with the customers and offering them freebies such as discounts on future service request in exchange for not opting out of their payment in order to reduce revenue loss.