

Background

Yellevate is a company that specializes in providing marketing services to other companies. It helps mid-sized companies to launch their marketing operations, which includes things like email marketing, website development, content creation, and others.

For the past few years, Yellevate has been struggling with client disputes. Yellevate defines disputes as clients expressing dissatisfaction with the company's service and refusing to pay for them.

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

Problem Statement

Yellevate's profit is heavily affected due to contract technicality and disputes between the company and clients

Objectives

This project aims to:

1. Identify the total profit gained and loss by Yellevate
 - a. Identify the overall number of dispute and not dispute
 - i. Among the disputed, how many are losses and wins
 - ii. Among the disputed, how many are losses and wins by Country
2. Identify date ranges for invoices settlement of the following:
 - a. From the disputed and undisputed
 - b. Based on the dispute, how many did Yellevate win and lose the dispute?
 - i. How much did Yellevate profit?
 - ii. How much did Yellevate lose?
3. To determine the project goals set by Yellevate
 - a. The Processing Time In Which Invoices Are Settled (average# of days rounded to a whole number)
 - b. The Processing Time For The Company To Settle Disputes(average# of days rounded to a whole number)
 - c. Percentage Of Revenue Lost From Disputes(within two decimal places)
 - d. Percentage of disputes received by the company that were lost(within two decimal places)
 - e. The country where the company reached highest losses from lost disputes (inUSD).
4. Provide findings and recommendations to Yellevate

Methodology

Data Cleaning using MS Excel

1. Download the report from Refocus Thinkific
2. Name column L as Days Range Settled, enter the formula below and drag the formula down to the last item

```
=IF(J2=0,"0 day",IF(AND(J2>=1,J2<=10),"1-10 days",IF(AND(J2>=11,J2<=20),"11-20 days",IF(AND(J2>=21,J2<=30),"21-30 days",IF(AND(J2>=31,J2<=40),"31-40 days",IF(AND(J2>=41,J2<=50),"41-50 days",IF(AND(J2>=51,J2<=60),"51-60 days",IF(AND(J2>=61,J2<=70),"61-70 days","more than 70 days"))))))))
```

3. Name column M as Days Range Late, enter the formula below and drag the formula down to the last item

```
=IF(K2=0,"0 day",IF(AND(K2>=1,K2<=10),"1-10 days",IF(AND(K2>=11,K2<=20),"11-20 days",IF(AND(K2>=21,K2<=30),"21-30 days",IF(AND(K2>=31,K2<=40),"31-40 days","more than 40 days")))))
```

4. Name column N as Disputed, enter the formula below and drag the formula down to the last item

```
=IF(G2=1,"Disputed","Not Disputed")
```

5. Name column O as Disputed (Loss/Win), enter the formula below and drag the formula down to the last item

```
=IF(AND(G2=0,H2=0),"No dispute",IF(AND(G2=1,H2=1),"Loss","Win"))
```

Data Cleaning Using SQL

1. Create a duplicate table for yellevate invoices.

Syntax in creating a new table

```
CREATE TABLE yellevate_invoices_copy  
AS SELECT  
*  
FROM yellevate_invoices;
```

```

1 SELECT
2 *
3 FROM yellevate_invoices_copy

```

	country	customer_ID	invoice_number	invoice_date	due_date	invoice_amount_usd	disputed	dispute_lost	settled_date	days_to_settle	days_late
1	Spain	9286-VLKMI	101747306	2021-05-27	2021-06-26	5819	0	0	2021-05-29	2	0
2	France	6833-ETVHD	103483331	2020-07-27	2020-08-26	7213	0	0	2020-08-18	22	0
3	China	6160-HCSFI	104628267	2020-01-13	2020-02-12	7272	0	0	2020-02-04	22	0
4	China	1080-NDGAE	106360977	2020-02-19	2020-03-20	9348	1	0	2020-03-26	36	6
5	Spain	6296-UKEUZ	106486147	2021-09-21	2021-10-21	4655	0	0	2021-09-25	4	0
6	United States	6831-FIODB	110122785	2020-10-23	2020-11-22	3258	0	0	2020-11-30	38	8
7	Spain	6296-UKEUZ	120364375	2021-04-23	2021-05-23	3463	0	0	2021-04-24	1	0
8	Russia	5148-SYKLB	121797094	2020-09-28	2020-10-28	8915	0	0	2020-10-30	32	2
9	China	0379-NFVHP	122662308	2021-08-05	2021-09-04	4364	0	0	2021-08-21	16	0

2. Determine the minimum, average, and maximum invoice amount in USD.

Syntax:

SELECT

MAX (invoice_amount_usd) AS max_invoice_amount_usd,

ROUND (AVG(invoice_amount_usd),0) AS avg_invoice_amount_usd,

MIN (invoice_amount_usd) AS min_invoice_amount_usd

FROM yellevate_invoices;

	max_invoice_amount_usd	avg_invoice_amount_usd	min_invoice_amount_usd
	numeric	numeric	numeric
1	12828	5990	526

3. Determine the total amount of invoice

SELECT

SUM(invoice_amount_usd) AS potential_profit

FROM yellevate_invoices;

	potential_earnings
	numeric
1	14770318

4. Add new column to convert the values "0" and "1" from the column disputed. Let 0 be no disputes with client while 1 would be disputes with client

Syntax:

```
ALTER TABLE yellevate_invoices_copy  
ADD COLUMN disputed_converted varchar;
```

```
UPDATE yellevate_invoices_copy  
SET disputed_converted = (CASE  
    WHEN disputed = '0' THEN 'No Dispute with Client'  
    WHEN disputed = '1' THEN 'Dispute with Client'  
END);
```

```
SELECT *  
FROM yellevate_invoices_copy;
```

Output:

Data output Messages Notifications											
	order_ID integer varying	invoice_number numeric	invoice_date date	due_date date	invoice_amount_usd numeric	disputed numeric	dispute_lost numeric	settled_date date	days_to_settle integer	days_late integer	disputed_converted character varying
1	5-VLKMI	101747306	2021-05-27	2021-06-26	5819	0	0	2021-05-29	2	0	No Dispute with Cli...
2	3-ETVHD	103483331	2020-07-27	2020-08-26	7213	0	0	2020-08-18	22	0	No Dispute with Cli...
3	0-HCSFI	104628267	2020-01-13	2020-02-12	7272	0	0	2020-02-04	22	0	No Dispute with Cli...
4	0-NDGAE	106360977	2020-02-19	2020-03-20	9348	1	0	2020-03-26	36	6	Dispute with Client
5	5-UKEUZ	106486147	2021-09-21	2021-10-21	4655	0	0	2021-09-25	4	0	No Dispute with Cli...
6	1-FIODB	110122785	2020-10-23	2020-11-22	3258	0	0	2020-11-30	38	8	No Dispute with Cli...
7	5-UKEUZ	120364375	2021-04-23	2021-05-23	3463	0	0	2021-04-24	1	0	No Dispute with Cli...
8	3-SYKLB	121797094	2020-09-28	2020-10-28	8915	0	0	2020-10-30	32	2	No Dispute with Cli...
9	3-NFVHP	122662308	2021-08-05	2021-09-04	4364	0	0	2021-08-21	16	0	No Dispute with Cli...

5. Add new column for the converted dispute lost. Let 0 be Yellevate winning the dispute whereas 1 Yellevate losing the dispute

```
ALTER TABLE yellevate_invoices_copy  
ADD COLUMN disputed_lost_converted varchar;
```

```
UPDATE yellevate_invoices_copy  
SET disputed_lost_converted = (CASE  
    WHEN disputed = '0' THEN 'Yellevate won Dispute'  
    WHEN disputed = '1' THEN 'Yellevate lost Dispute'  
END);
```

```
SELECT *  
FROM yellevate_invoices_copy;
```

Output:

disputed numeric	dispute_lost numeric	settled_date date	days_to_settle integer	days_late integer	disputed_converted character varying	disputed_lost_converted character varying
0	0	2021-05-29	2	0	No Dispute with Cli...	Yellevate Won Dispute
0	0	2020-08-18	22	0	No Dispute with Cli...	Yellevate Won Dispute
0	0	2020-02-04	22	0	No Dispute with Cli...	Yellevate Won Dispute
1	0	2020-03-26	36	6	Dispute with Client	Yellevate Won Dispute
0	0	2021-09-25	4	0	No Dispute with Cli...	Yellevate Won Dispute
0	0	2020-11-30	38	8	No Dispute with Cli...	Yellevate Won Dispute
0	0	2021-04-24	1	0	No Dispute with Cli...	Yellevate Won Dispute
0	0	2020-10-30	32	2	No Dispute with Cli...	Yellevate Won Dispute
0	0	2021-08-21	16	0	No Dispute with Cli...	Yellevate Won Dispute

6. Determine how many are at disputes with Yellevate

```
SELECT disputed_converted FROM yellevate_invoices_copy;
```

```
SELECT
COUNT (CASE
  WHEN "disputed_converted" ILIKE 'Dispute with Client' THEN 1
  ELSE NULL
END) as Count_dispute,
COUNT (CASE
  WHEN "disputed_converted" ILIKE '%No Dispute with Client%' THEN 1
  ELSE NULL
END) AS count_no_dispute
FROM yellevate_invoices_copy
```

Output:

Data output	Messages	Notifications
<div> <div>+</div> <div>📄</div> <div>▼</div> <div>📋</div> <div>🗑️</div> <div>🔄</div> <div>⬇️</div> </div>		
	count_dispute bigint	count_no_dispute bigint
1	571	1895

7. Count of disputes have already be determined, next, which countries have the highest dispute?

Syntax:

```
SELECT
country,
COUNT (CASE
  WHEN "disputed_converted" ILIKE 'Dispute with Client' THEN 1
  ELSE NULL
END) as Count_dispute,
COUNT (CASE
  WHEN "disputed_converted" ILIKE '%No Dispute with Client%' THEN 1
  ELSE NULL
```

```

END) AS count_no_dispute
FROM yellevate_invoices_copy
GROUP BY country
ORDER BY country ASC

```

Data output Messages Notifications			
	country character varying	count_dispute bigint	count_no_dispute bigint
1	China	61	555
2	France	222	339
3	Russia	149	238
4	Spain	59	337
5	United States	80	426

8. From those who were at disputes with Yellevate, how many did Yellevate win? How many were losses?

Syntax:

SELECT

```




COUNT (CASE
  WHEN "disputed_lost_converted" ILIKE '%won%' THEN 1
  ELSE NULL
END) as yellevate_won,
COUNT (CASE
  WHEN "disputed_lost_converted" ILIKE '%lost%' THEN 1
  ELSE NULL
END) AS yellevate_lost
FROM yellevate_invoices_copy

```

Data output Messages Notification		
	yellevate_won bigint	yellevate_lost bigint
1	2365	101

9. Based on the answer from number 8, determine now the countries who have pursued to have disputes with Yellevate

```
SELECT
Country,
COUNT (CASE
    WHEN "disputed_lost_converted" ILIKE '%won%' THEN 1
    ELSE NULL
    END) as yellevate_won,
COUNT (CASE
    WHEN "disputed_lost_converted" ILIKE '%lost%' THEN 1
    ELSE NULL
    END) AS yellevate_lost
FROM yellevate_invoices_copy
GROUP BY country
ORDER BY country ASC
```

	country character varying 	yellevate_won bigint 	yellevate_lost bigint 
1	China	611	5
2	France	485	76
3	Russia	374	13
4	Spain	392	4
5	United States	503	3

10. Create a days range for days settled

```
ALTER TABLE yellevate_invoices_copy
ADD COLUMN days_settled_date_range varchar;

UPDATE yellevate_invoices_copy
SET days_settled_date_range = (CASE
    WHEN days_to_settle = '0' THEN '0 day'
    WHEN days_to_settle <= '10' THEN '1-10 Days'
    WHEN days_to_settle <= '20' THEN '11-20 Days'
    WHEN days_to_settle <= '30' THEN '21-30 Days'
    WHEN days_to_settle <= '40' THEN '31-40 Days'
    WHEN days_to_settle <= '50' THEN '41-50 Days'
    WHEN days_to_settle <= '60' THEN '51-60 Days'
    WHEN days_to_settle <= '70' THEN '61-70 Days'
    WHEN days_to_settle <= '80' THEN '71-80 Days'
    ELSE NULL
    END);
```

```
UPDATE 2466
```

```
Query returned successfully in 134 msec.
```

```
SELECT
days_to_settle, days_settled_date_range
FROM yellevate_invoices_copy;
```

Data output Messages Notifications		
	days_to_settle integer	days_settled_date_range character varying
1	18	11-20 Days
2	49	41-50 Days
3	23	21-30 Days
4	31	31-40 Days
5	22	21-30 Days
6	14	11-20 Days
7	38	31-40 Days
8	41	41-50 Days
9	9	1-10 Days
10	31	31-40 Days
11	32	31-40 Days
12	32	31-40 Days
13	28	21-30 Days
14	37	31-40 Days
15	37	31-40 Days

11. After accomplishing no. 10, filter it by country






```
SELECT
country,
COUNT (CASE
    WHEN "days_settled_date_range" ILIKE '%Zero%' THEN 1
    ELSE NULL
END) as O0day,
COUNT (CASE
    WHEN "days_settled_date_range" ILIKE '%1-10%' THEN 1
    ELSE NULL
END) as O1_to_10_days,
```








```

COUNT (CASE
  WHEN "days_settled_date_range" ILIKE '%11-20%' THEN 1
  ELSE NULL
END) as O11_to_20_days,
COUNT (CASE
  WHEN "days_settled_date_range" ILIKE '%21-30%' THEN 1
  ELSE NULL
END) as O21_to_30_days,
COUNT (CASE
  WHEN "days_settled_date_range" ILIKE '%31-40%' THEN 1
  ELSE NULL
END) as O31_to_40_days,
COUNT (CASE
  WHEN "days_settled_date_range" ILIKE '%41-50%' THEN 1
  ELSE NULL
END) as O41_to_50_days,
COUNT (CASE
  WHEN "days_settled_date_range" ILIKE '%51-60%' THEN 1
  ELSE NULL
END) as O51_to_60_days,
COUNT (CASE
  WHEN "days_settled_date_range" ILIKE '%61-70%' THEN 1
  ELSE NULL
END) as O61_to_70_days,
COUNT (CASE
  WHEN "days_settled_date_range" ILIKE '%71-80%' THEN 1
  ELSE NULL
END) as O71_to_80_days
FROM yellevate_invoices_copy
GROUP BY country
;

```

	country character varying 	o0day bigint 	o1_to_10_days bigint 	o11_to_20_days bigint 	o21_to_30_days bigint 
1	Russia	2	30	73	122
2	China	0	70	174	215
3	Spain	2	62	89	112
4	France	0	52	119	157
5	United States	0	21	107	182

o31_to_40_days bigint 	o41_to_50_days bigint 	o51_to_60_days bigint 	o61_to_70_days bigint 	o71_to_80_days bigint 
78	58	22	2	0
115	39	3	0	0
75	41	13	2	0
145	70	15	2	1
126	49	20	1	0

12. Create days range for days late

```
UPDATE yellevate_invoices_copy
SET days_late_converted = (CASE
  WHEN days_late = '0' THEN '0 Day'
  WHEN days_late <= '10' THEN '1-10 Days'
  WHEN days_late <= '20' THEN '11-20 Days'
  WHEN days_late <= '30' THEN '21-30 Days'
  WHEN days_late <= '40' THEN '31-40 Days'
  WHEN days_late <= '50' THEN '41-50 Days'
  WHEN days_late <= '60' THEN '51-60 Days'
  WHEN days_late <= '70' THEN '61-70 Days'
  WHEN days_late <= '80' THEN '71-80 Days'
  ELSE NULL
END);
```

13. The Processing Time In Which Invoices Are Settled (average# of days rounded to a whole number)

```
SELECT
ROUND(AVG(days_to_settle),0) AS avg_days_to_settle
FROM yellevate_invoices_copy
```

	avg_days_to_settle	
	numeric	
1		26

14. The Processing Time For The Company To Settle Disputes(average# of days rounded to a whole number)


```
SELECT
ROUND(AVG(days_to_settle),0) AS avg_dispute_loss
FROM yellevate_invoices_copy
WHERE dispute_lost = 1;
```

	avg_dispute_loss	
	numeric	
1		34

15. Percentage Of Revenue Lost From Disputes(within two decimal places)


```
SELECT
```

```
ROUND(COUNT(dispute_lost) *100.0 / (SELECT COUNT(dispute_lost) FROM
yellevate_invoices_copy),2)
FROM yellevate_invoices_copy
WHERE dispute_lost=1
```

	round numeric 
1	4.10

16. Percentage of disputes received by the company that were lost(within two decimal places)

```
SELECT
ROUND(SUM(invoice_amount_usd) *100.0 /
(SELECT SUM(invoice_amount_usd)
FROM yellevate_invoices_copy
),2)
FROM yellevate_invoices_copy
WHERE dispute_lost=1;
```

	round numeric 
1	4.67

17. The country where the company reached highest losses from lost disputes (inUSD).

```
SELECT
country,
COUNT (CASE
WHEN "disputed_lost_converted" ILIKE '%lost%' THEN 1
ELSE NULL
END) AS yellevate_lost ,
SUM (invoice_amount_usd) AS total_loss_disputes
FROM yellevate_invoices_copy
WHERE dispute_lost = 1
GROUP BY country
ORDER BY total_loss_disputes DESC;
```

	country character varying 🔒	year_voted_lost bigint 🔒	total_loss_disputes numeric 🔒
1	France	76	526264
2	Russia	13	81291
3	China	5	42630
4	United States	3	22936
5	Spain	4	17046

Findings

Get the Data Analysis Goals

1. The processing time in which invoices are settled

Ave. days the invoice are settled
26

Formula used: =AVERAGE('RAW FILE'!J2:J2467)

2. The processing time for the company to settle dispute

Ave. days company settled disputes
34

Formula used: =AVERAGEIF('RAW FILE'!O:O,"Loss",'RAW FILE'!J:J)

3. Percentage of disputes received by the company that were lost (within two decimal places)

Dispute/Not Disputed	% of Dispute and Not Disputed
Not disputed	76.85%
Not disputed	76.85%
Disputed	23.15%
Win	19.06%
Loss	4.10%
Grand Total	100.00%

Used pivot, count all the dispute (loss) and divided by the grand total

4. Percentage of revenue lost disputes (within two decimal places)

Dispute/Not Disputed	Revenue % of Dispute and Not Disputed
Not disputed	74.62%
Not disputed	74.62%
Disputed	25.38%
Loss	4.67%
Win	20.71%
Grand Total	100.00%

Used pivot, sum the usd amount all the dispute (loss) and divided by the grand total

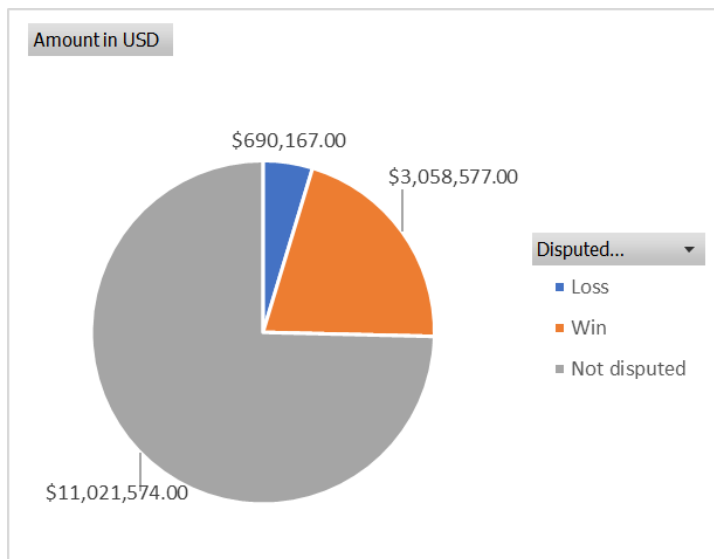
5. The country where the company reached the highest losses from the lost disputes (in USD)

Country	Revenue Loss
France	\$526,264.00
Russia	\$81,291.00
China	\$42,630.00
United States	\$22,936.00
Spain	\$17,046.00
Grand Total	\$690,167.00

Use pivot, Country and sort by highest loss

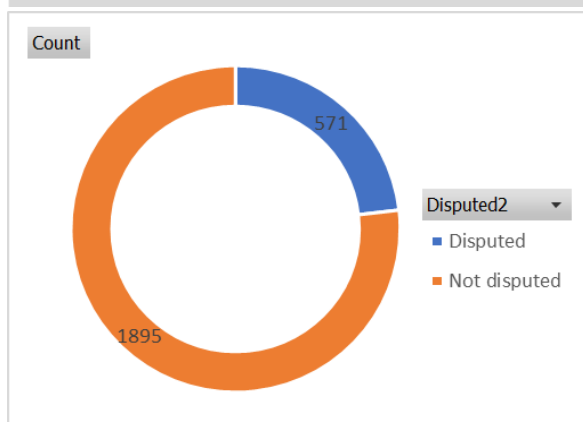
There is a \$14,770,318.00 total amount of revenue by Yellevate
Dispute - Loss amounting to 690,167.00 is 5% of the total Yellevate revenue

Loss/Win/Not Disputed	Amount in USD
Loss	\$690,167.00
Win	\$3,058,577.00
Not disputed	\$11,021,574.00
Grand Total	\$14,770,318.00

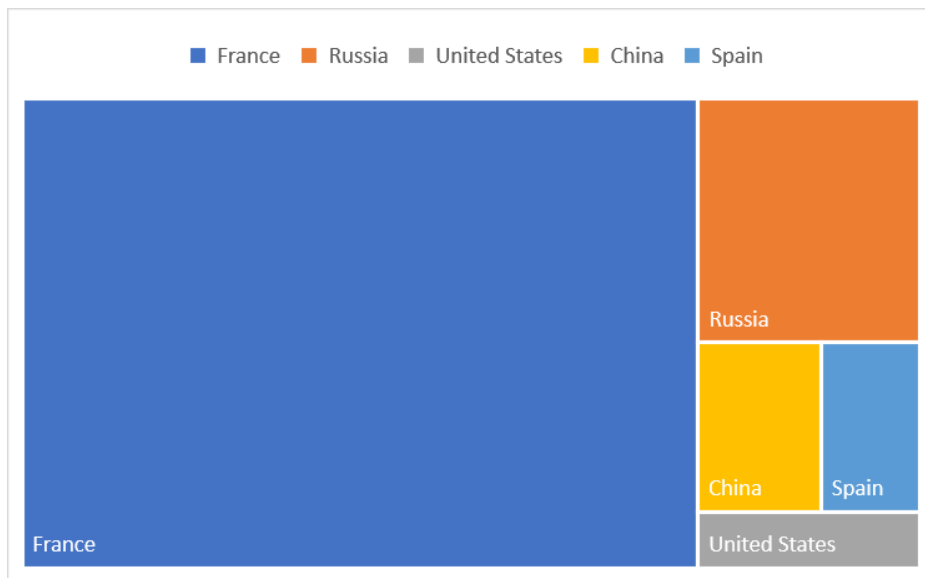


Out of 2466 invoices, 30% (571) has been disputed by the client.

Disputed/Not Disputed	Count
Disputed	571
Not disputed	1895
Grand Total	2466

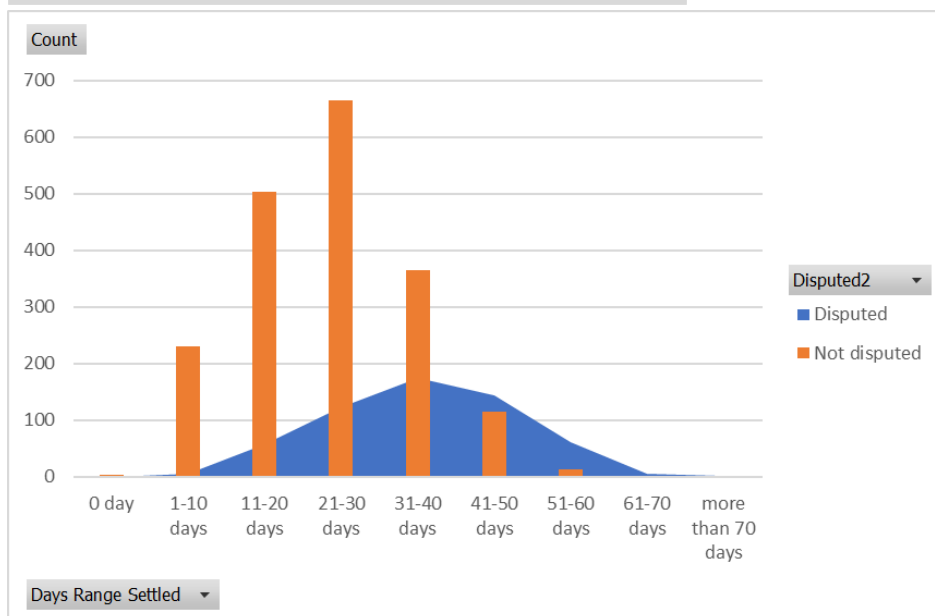


France was the highest contributor to the dispute. Which is the 39% of all disputes



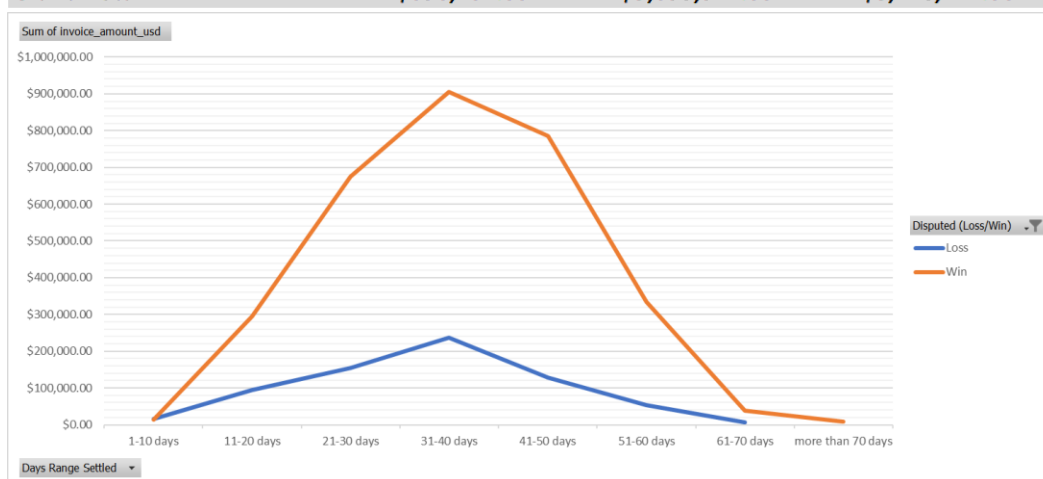
Those invoices that was settled on 31-40 days has the highest rate of dispute

Settlement Days Ranges	Disputed	Not disputed
0 day		4
1-10 days	5	230
11-20 days	58	504
21-30 days	123	665
31-40 days	174	365
41-50 days	143	114
51-60 days	61	12
61-70 days	6	1
more than 70 days	1	
Grand Total	571	1895



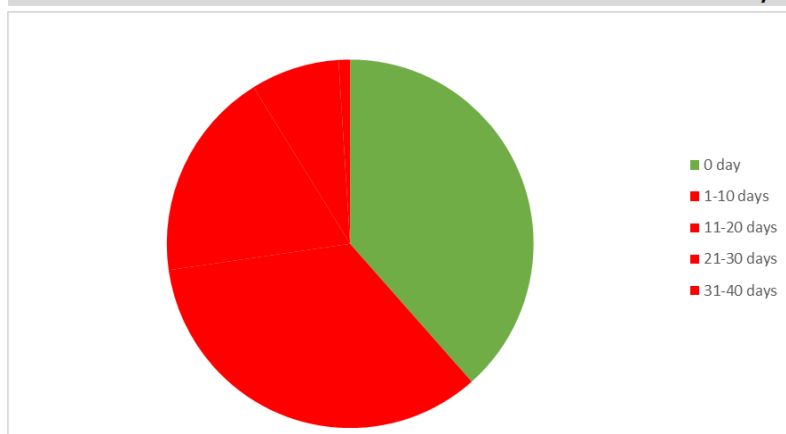
The peak of dispute loss and win happens after 30 days the standard contract due date. Those invoices that was settled on 31-40 days has the highest count of dispute win and loss

Settlement Days Ranges	Loss	Win	Grand Total
1-10 days	\$16,790.00	\$14,694.00	\$31,484.00
11-20 days	\$94,348.00	\$295,156.00	\$389,504.00
21-30 days	\$154,301.00	\$675,720.00	\$830,021.00
31-40 days	\$236,393.00	\$906,121.00	\$1,142,514.00
41-50 days	\$127,421.00	\$785,469.00	\$912,890.00
51-60 days	\$53,919.00	\$334,063.00	\$387,982.00
61-70 days	\$6,995.00	\$38,715.00	\$45,710.00
more than 70 days		\$8,639.00	\$8,639.00
Grand Total	\$690,167.00	\$3,058,577.00	\$3,748,744.00



Out of those dispute loses, 61.54% happens after 30 days which cost USD 424,728 out of USD 690,167

Days Late Ranges	Loss	Grand Total
0 day	38.46%	38.46%
1-10 days	34.25%	34.25%
11-20 days	18.46%	18.46%
21-30 days	7.81%	7.81%
31-40 days	1.01%	1.01%
Grand Total	100.00%	Days Late Ranges



Recommendation

1. Yellevate should reach out to the top 10 clients who always dispute the service provided to them. In this way, Yellevate could have an idea on the reason for the dispute and provide a solution if necessary. The disputes of the top 10 is 38% of overall disputes

Rank	Customer ID	Count of Disputed2
1	3448-OWJOT	27
2	3568-JJMFV	25
3	1080-NDGAE	24
4	9725-EZTEJ	24
5	8102-ABPKQ	23
6	7600-OISKG	22
7	9771-QTLGZ	21
8	8389-TCXFQ	17
9	6048-QPZCF	17
10	4632-QZOKX	17

2. In order to recover some of the loss of Yellevate a rule in the contract must be created. See below

Day Ranges Settlement	Remarks
1-20 days	Client can dispute
21-25 days	Decision should be made on dispute
26-30 days	Window Yellevate and Client to settle
31-40 days	Yellevate and Client to be penalize of 10% of invoice
41-45 days	Yellevate and Client to be penalize of 15% of invoice
More than 46 days	Yellevate and Client to be penalize of 20% of invoice

Win/Loss	Sum of invoice_amount_usd	Sum of Amount with Penalty	Sum of Difference
Loss	\$690,167.00	\$632,363.95	-\$57,803.05
1-10 days	\$16,790.00	\$16,790.00	\$0.00
11-20 days	\$94,348.00	\$94,348.00	\$0.00
21-30 days	\$154,301.00	\$154,301.00	\$0.00
31-40 days	\$236,393.00	\$212,753.70	-\$23,639.30
41-50 days	\$127,421.00	\$105,440.05	-\$21,980.95
51-60 days	\$53,919.00	\$43,135.20	-\$10,783.80
61-70 days	\$6,995.00	\$5,596.00	-\$1,399.00
Win	\$3,058,577.00	\$3,358,919.60	\$300,342.60
1-10 days	\$14,694.00	\$14,694.00	\$0.00
11-20 days	\$295,156.00	\$295,156.00	\$0.00
21-30 days	\$675,720.00	\$675,720.00	\$0.00
31-40 days	\$906,121.00	\$996,733.10	\$90,612.10
41-50 days	\$785,469.00	\$918,916.10	\$133,447.10
51-60 days	\$334,063.00	\$400,875.60	\$66,812.60
61-70 days	\$38,715.00	\$46,458.00	\$7,743.00
more than 70 days	\$8,639.00	\$10,366.80	\$1,727.80
Grand Total	\$3,748,744.00	\$3,991,283.55	\$242,539.55

For example we used the above rules on the current data.

You can see above on Loss portion, Yellevate will less 57,803.05 which make the total loss revenue to become 632,363.95 from 690,167.00 However, it gained 300,342.60 additional to the win amount 3,058,577.00 for the total win revenue of 3,358,919.60. The total profit will become 15,012,857.55 from 14,770,318.00. Which gained Yellevate at least 242,539.55 which is 1.6% of the previous profit. Total loss percentage has now reduced to 4% from 5%