Bolt

Business Case Study and Report

Presented by: Guilherme Luz. 🇧🇷

Bolt

Problem Statement

- Reduce useless time spent by a courier
- Understanding why orders are failing
- Develop business approaches based on the discoveries

Challenges and Objectives

Challenges

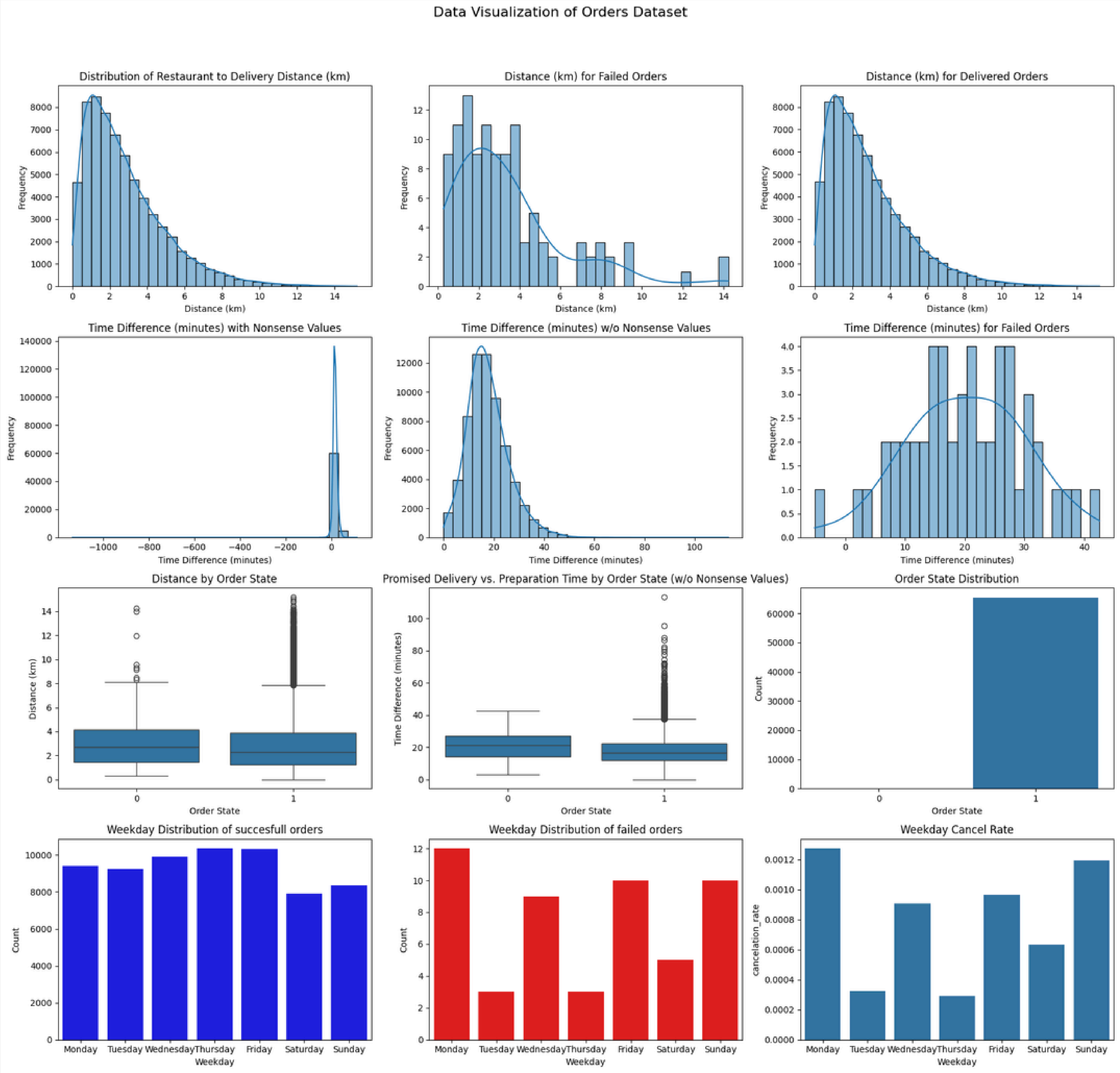
- Data analysis of the Orders and Order Stages datasets.
- Feature engineering to create meaningful variables.
- Investigating why orders are canceled and identifying the most impactful factors.
- Determining whether failed orders are due to courier issues, timing problems, or other factors.
- Developing a structured methodology to assess delivery performance and user experience.
- Designing data models and identifying viable features to address the business problem effectively.

Methodology

	DESCRIPTION	INSIGHT
#1. Feature engineering and data visualization of orders dataset	Create features for distance using latitude and longitude, incorporate weekdays, and visualize the data to extract meaningful insights.	1. Successful orders were more likely to have a shorter distance and a shorter time difference between promised delivery and finished preparation. 2. Weekdays are relevant in order sucess
#2. Feature engineering and merging dataset of order stages	Creating a feature of restaurant timing, preparing data for statisticall analysis	1. Hypothesis: Canceled orders have longer promised delivery times. 2. Hypothesis: Vehicle type may influence the delivery experience. 3. Hypothesis: Canceled orders may have longer time intervals between the beginning and completion of orders compared to completed ones. 4. Hypothesis: Distance may influence order cancellations.

	DESCRIPTION	INSIGHT
#3. Evaluating each hypothesis individually	Create new features for time-related variables, such as dif_finished_expected (the difference between restaurant finished food preparation and delivery expectation - also known as courier delivery time)	1. The features could bring light to the hypothesis that time is the key feature for delivery success
#4. Statistical Analysis	Perform a series of statistical tests to evaluate weather the hypothesis are accepted or not (ANOVA, Chi-Squared, t-test, etc..)	1. Only time-related variables were actually statistically different between delivered and failed orders
#5. Machine Learning model for order state prediction	Could we try to predict order success based on given features? Well, I tryed to develop a way to predict wheather a specific order could be delivered or not, that could help us to develop future app features and comercial applications.	1. The model was not successfull in predict order state due to lack of data for failed orders. We actually have very little information about that order, less than 0.0001% of all orders were failed in this dataset.
#6. Machine Learning model for predicting warm food arrival	Could we try to predict wheather a delivery would arrive in 20 minutes or less? This is an assumption that I made for the order to arrive warm and tasty.😋	1. That model was particularlly successfull, using only restaurant distance, vehicle type and two time-related variables. 2. We are now able to predict 20 minutes delivery with 90% precision.

First data exploration



Insights

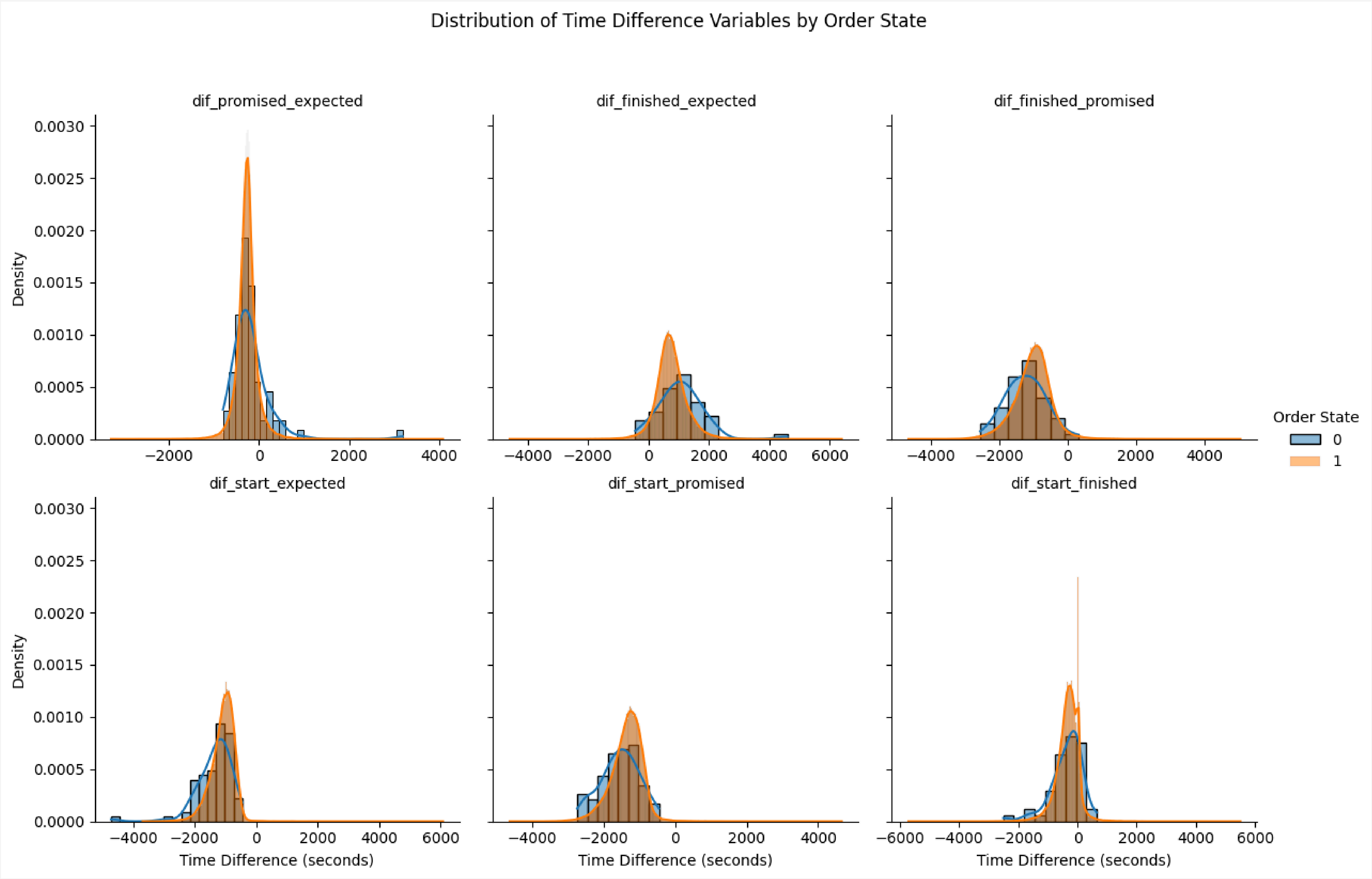
- Generally, orders that were sucesfully delivered were more likely to have a shorter distance and a shorter time difference between promised delivery and preparation.

- The amount of failed orders were not very significant in the dataset.

- There are also some outliers in the dataset that should be investigated.

- Clearlly we have more cases of a cancelation on Monday and Sunday that we must keep attention to.

Feature creating in time



Feature listing:

dif_promised_expected:
Measures the gap between when the order was promised to be delivered and our updated delivery estimate.

dif_finished_expected
Measures the time difference between when the restaurant finished preparing the order and the updated delivery estimate.

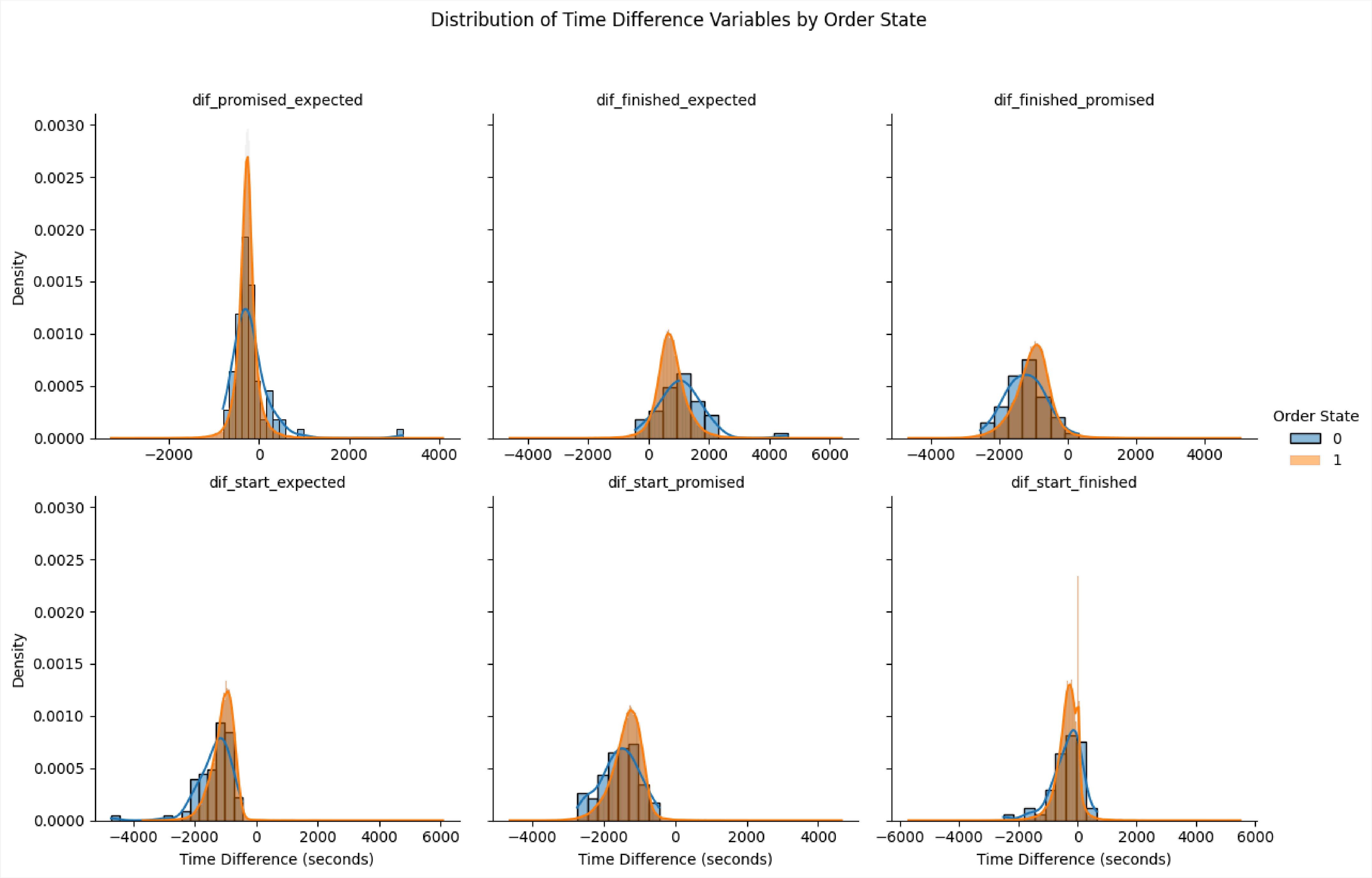
dif_finished_promised
Measures the time difference between when the restaurant finished preparing the order and the promised delivery time.

dif_start_expected
Measures the time difference between when the order was requested by the user and the updated delivery estimate.

dif_start_promised
Measures the time difference between when the order was requested by the user and the promised delivery time.

dif_start_finished
Measures the time difference between when the order was requested by the user and when the restaurant finished preparing the order.

Feature creating in time



Now we can see more clearly that non-delivered orders have a slight difference in the distribution of the time difference variables.

- The difference between finished preparation and updated expected delivery is higher for non-delivered orders, this might indicate that the user may feel frustrated with the delivery time.
- The same behavior as above can be seen in finished preparation and promised delivery time.
- The difference between order started and updated delivery time is higher for non-delivery orders.
- Similar behavior as above for promised delivery and order started.

Feature creating in time

Column	Delivered_Normality_p	Not_Delivered_Normality_p	Test_Type	p_value
dif_promised_expected	0.0	1.044151e-12	Mann-Whitney U Test	0.803404
dif_finished_expected	0.0	2.366163e-04	Mann-Whitney U Test	0.000312
dif_finished_promised	0.0	9.995574e-01	Mann-Whitney U Test	0.004980
dif_start_expected	0.0	2.345167e-08	Mann-Whitney U Test	0.000002
dif_start_promised	0.0	4.088987e-01	Mann-Whitney U Test	0.000369
dif_start_finished	0.0	8.686852e-05	Mann-Whitney U Test	0.751094
dif_rest_time_avg	0.0	9.864240e-03	Mann-Whitney U Test	0.536236

expected - promised
transportation time
promised transportation time
from courier propose to delivery expectation
from courier propose to delivery promise
whole time delivery journey
variation of restaurant preparation time

The new variables: dif_finished_expected, dif_finished_promised, dif_start_expected, dif_start_promised actually had statisticall significance. ($p<0.05$)

It proves that those variables indeed showed differences between delivered and failed orders.

Time may be trully a key factor for order failure.

**This thrilling result does not repeats for other hypothesis (distance and courier performance)*

Order state predictor

```
Classification Report:
              precision    recall  f1-score   support

           0       0.00      0.00      0.00        24
           1       1.00      1.00      1.00     19232

    accuracy                           1.00     19256
   macro avg       0.50      0.50      0.50     19256
weighted avg       1.00      1.00      1.00     19256

ROC-AUC Score: 0.5120751438574598
```

	order_state	count_rows
0	0	79
1	1	64105

I developed this machine learning predictor of order state (failed or success), but due to imbalanced target features, I could not reach a good model.
It is not better than flipping a coin.
Needed more data to improve it.

20 min predictor

```
Classification Report:
              precision    recall  f1-score   support

       False       0.81      0.62      0.70      1070
        True       0.98      0.99      0.98     18171

    accuracy                           0.97     19241
   macro avg       0.89      0.81      0.84     19241
weighted avg       0.97      0.97      0.97     19241

ROC-AUC Score: 0.9803554446671471
```

I made an assumption that it takes around 20 minutes delivery for your food to arrive warm.
So I created a surprisingly good model to predict if your food will arrive in 20 minutes.

I used the same created variables to predict it.

But, well, what could we do with it?

Next steps

Bolt

NEXT STEPS

Improve failed orders amount of data

Add more features

Order related features:
1. Cousine type
2. Order value
3. More order stages

User related features:
1. Maturity level
2. Churn rates
3. Consume behavior

Develop improved Machine Learning Models based on it

Better undertand restaurant preparation time.

- I didn't have variable telling me exactly when order started, only the moment it was proposed to courier.

Scale 20 minutes models-like

Scalling

"If you deliver in less than 20 minutes you'll recieve extra compensation"

A/B test if it really has potential do scale

Hope to see you soon!