

FedEx Shipments

Lachandra Ash

November 10, 2022

Topic

Each of the FedEx subsidiaries is under the wing of FedEx Corporation. FedEx Express delivers packages to 220+ nations and regions, including every street address located in the U.S. FedEx Services offers critical back-office services to our transportation businesses, including sales and marketing, information technology, communications, customer service, technical support, billing, dand collection, and more (FedEx). FedEx Ground offers reliable business-to-business delivery in addition to simple residential service through FedEx Home Delivery® and FedEx Ground® Economy (FedEx).

Business Problem

FedEx want their revenues to grow faster, and they want their consumers to be satisfied with their services. FedEx consumers may be upset if their shipments are delayed. Delayed shipments can cause FedEx to lose profit if their consumers leave them permanently.

Background | History

FedEx's corporation was founded by Fredrick W. Smith. Fredrick W. Smith was a Yale graduate student, and he proposed a plan to create a shipment company (FedEx, n.d.). FedEx Ex was originally spelled out as Federal Express. Mr. Smith create the Federal Express name by thinking about the Federal Reserve Bank, the word federal, and economics (FedEx, n.d.). The Federal Express shipment operations began in Memphis, Tennessee. The first few packages were delivered to various cities on April 17, 1973 (FedEx, n.d.). The first FedEx drop box was installed in 1975.

FedEx Express, FedEx Ground, FedEx Custom Critical, FedEx Global Logistics, and FedEx Freight are all names of companies that were once acquired and relaunched as part of the FedEx family (FedEx, n.d.). FedEx Corporate Service, Inc., or FedEx Services. was developed to consolidate the company's branding, service and support, and information technology infrastructure for FedEx Express and FedEx Ground (FedEx, n.d.).

Data Explanation

The FedEx dataset has fifteen columns and 3,604,175 rows. The target variable is Delivery_Status, and the task type is classification. There are eight integer features, four float features, and three objects features. The dataset features are:

- Year: The Year the data was collected
- Month: The Month in which the data was collected
- DayofMonth: The day of the month
- DayofWeek: The day of Week
- ActualShipmentTime: The Actual time when the package was sent for shipment.
(ex:1955 means 19 hours and 55 minutes i.e 7:55 PM)
- PlannedShipmentTime: The time when the package should have been sent for shipment.
(ex: 1955 means 19 hours and 55 minutes i.e., 7:55 PM)

- **PlannedDeliveryTime:** The time when the package should be delivered. (ex: 1955 means 19 hours and 55 minutes i.e 7:55 PM)
- **CarrierName:** The name of the Carrier which carried the package.
- **CarrierNum:** The number of the Carrier which carried the package.
- **PlannedTimeofTravel:** The estimated time to reach from Source to Destination. (in minutes)
- **ShipmentDelay:** The time by which the package was shipped late. (in minutes. Negative value indicates that the package was shipped early. Ex: 4 indicates that the package was shipped 4 minutes late, whereas -4 indicates that the package was shipped 4 minutes early)
- **Source:** The place from which the package was shipped.
- **Destination:** The place at which the package was delivered.
- **Distance:** Distance between Source and Destination in miles.
- **Delivery_Status:** Delivered at right time or not. (Dependent Variable) (Ahmed, n.d.)

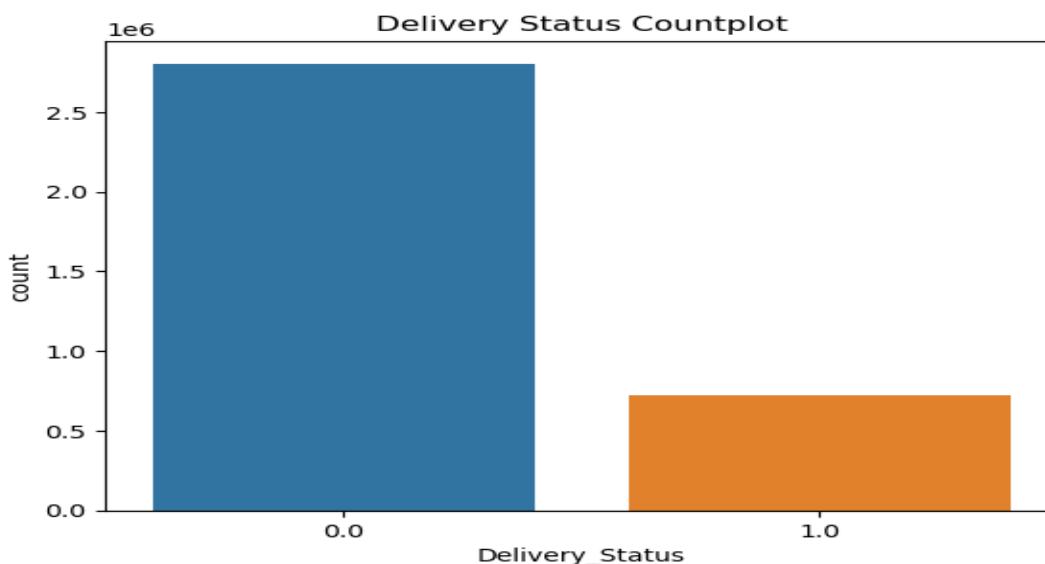
Methods

I obtained the FedEx dataset from the Kaggle website, and I used Jupyter Python Notebook 3 to analyze the data. I imported the necessary libraries, modules into the notebook, and imported the FedEx dataset into the pandas dataframe.

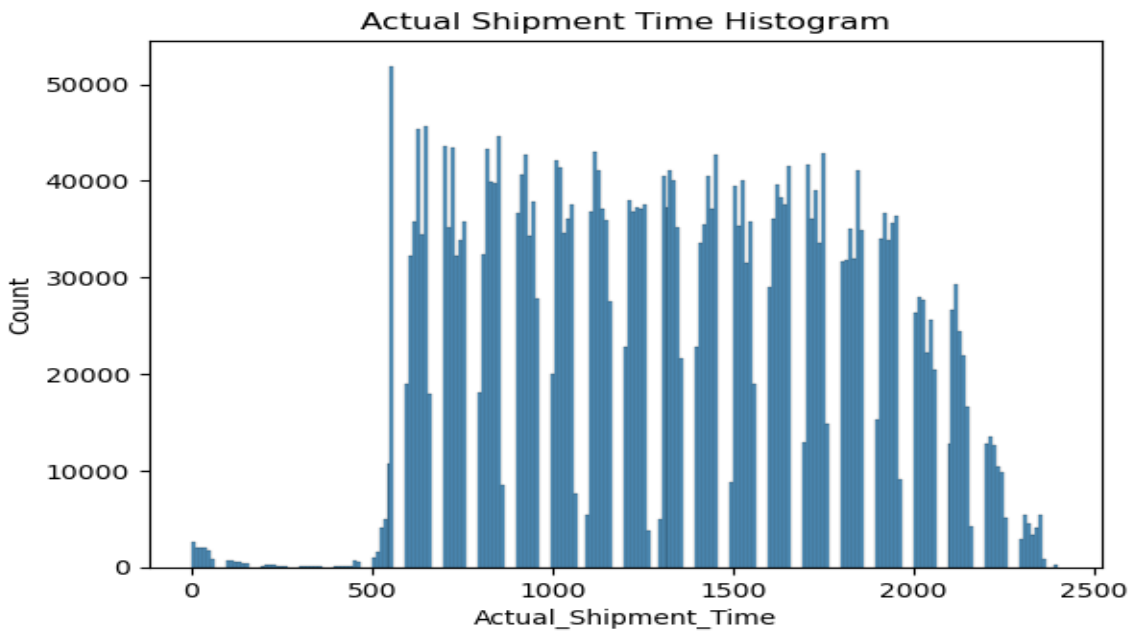
I cleansed the dataset by checking for missing values, checking the sum of missing values, filling the missing values with fillna method, checking delivery status records, and dropping destination, source, and carrier name columns. I examined the data by checking the index, dtypes, info, shape, describe, and columns.

Analysis

1) **Delivery Status:** There are more deliveries that were not delivered on time versus on time deliveries.



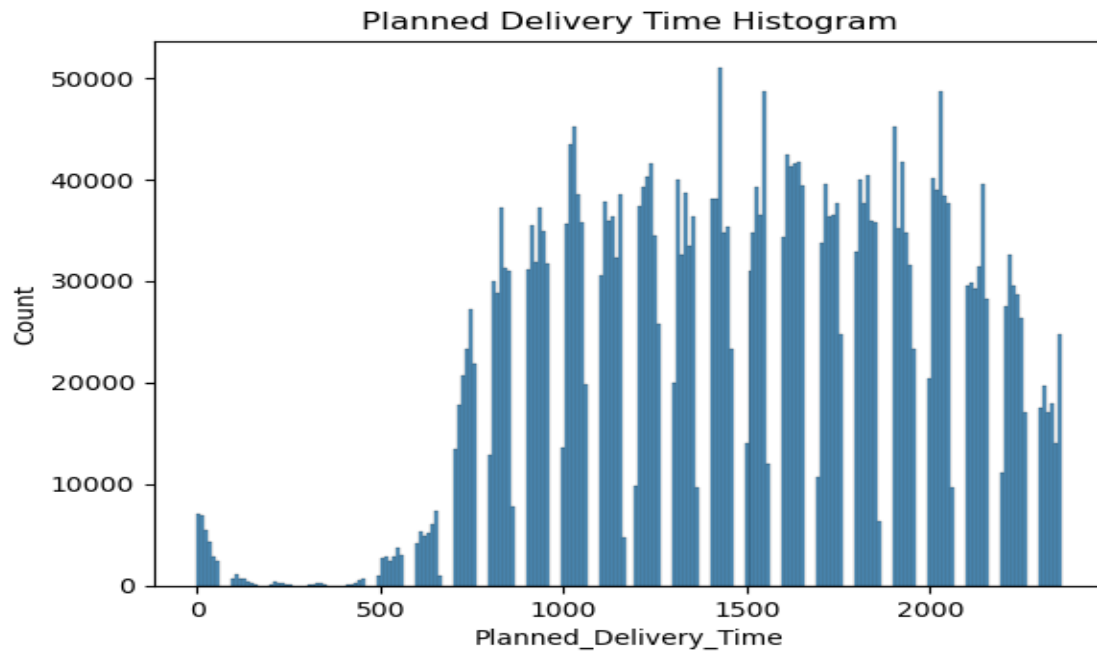
2)Actual Shipment Time: The actual shipment time began after 5am and ends after 8pm.



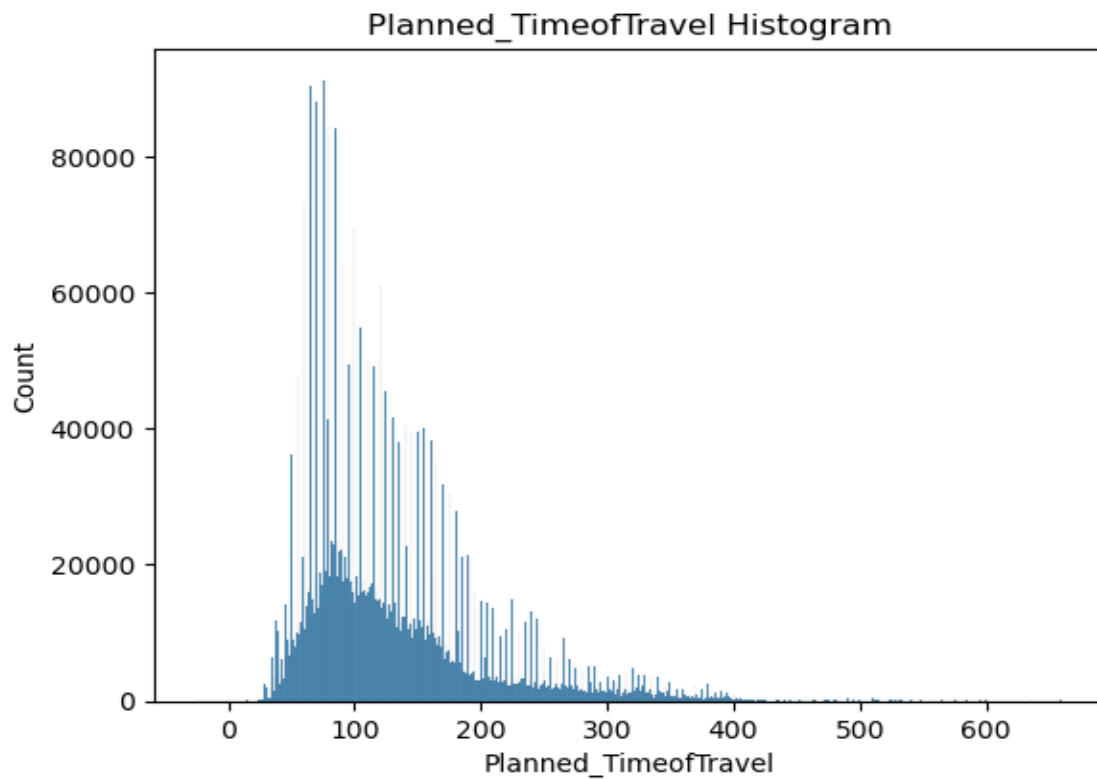
3)Planned Shipment Time Histogram: The actual shipment time began after 5am and ends after 8pm.



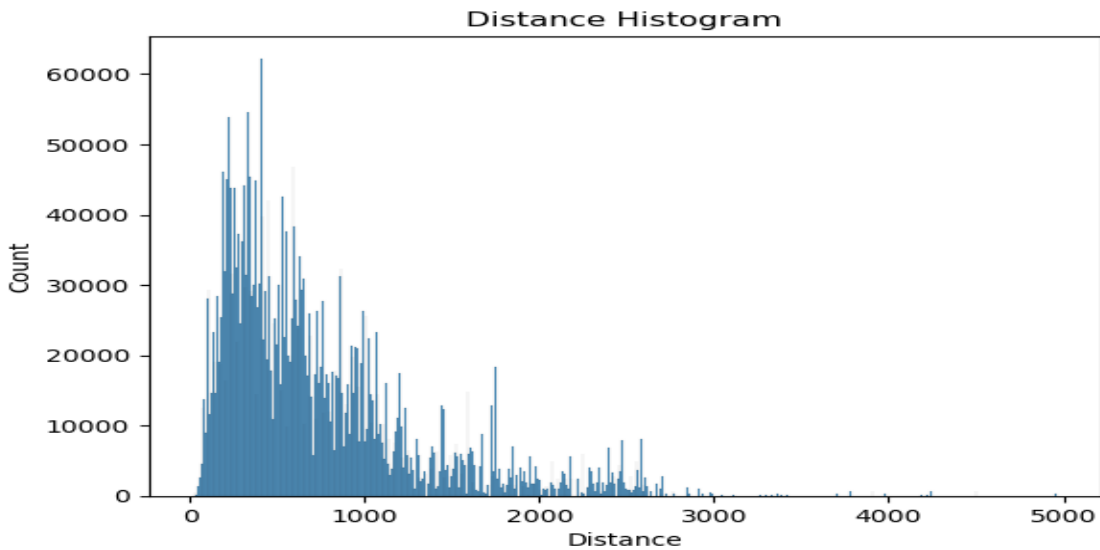
4)Planned Delivery Time: The actual shipment time began after 5am and ends after 8pm



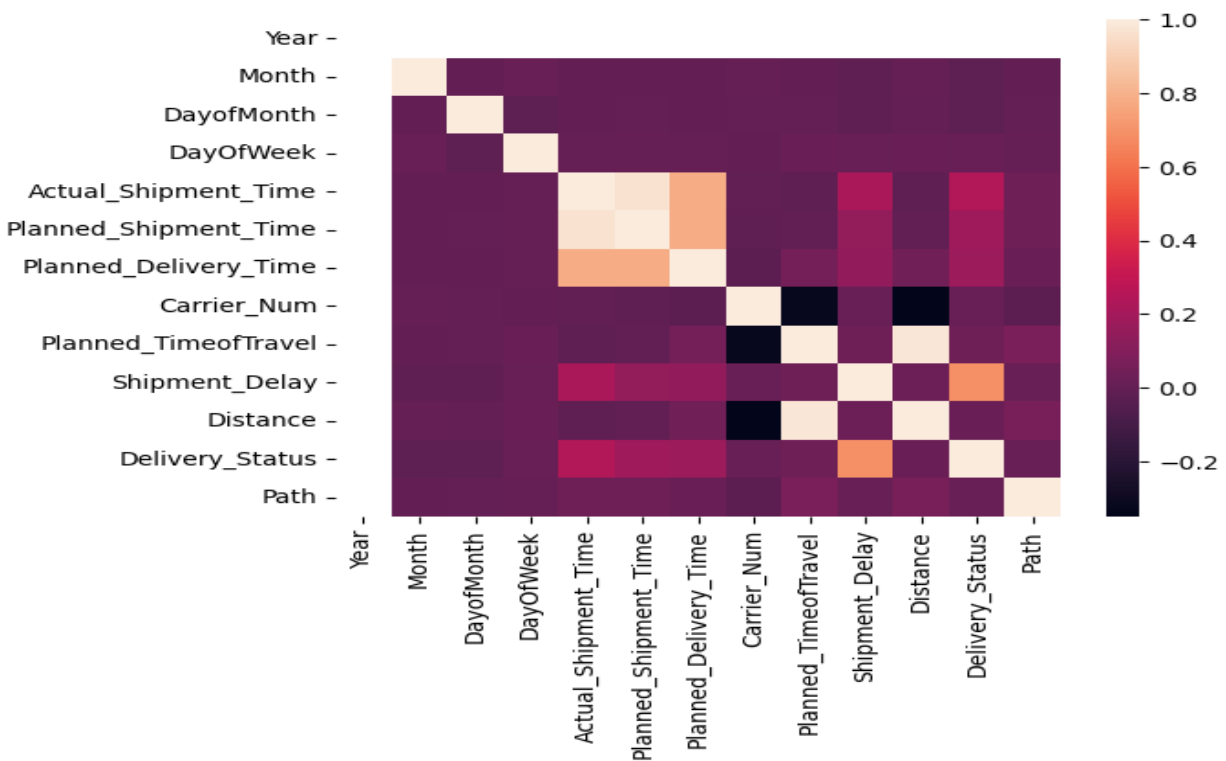
5)Planned time of travel: The shipments traveled for four hours till it reached its destination.



6)Distance: The greatest number of shipments traveled approximately between 0 to 500 miles. After 500 miles, the number of shipments began to decrease.



7)Heatmap Correlation: The heatmap displays the correlation between each feature within the dataframe.



The Model

I chose to use the Naïve Bayes model for my prediction score. I used the sklearn for train, test, split, GaussianNB, and to import and print the metrics. I used the accuracy score, classification report, and confusion matrix metrics. The model's accuracy score displayed 97.9%.

View Appendix, Table 1: Model Accuracy Scores

Conclusion

The precision score was 99% accuracy for shipments not being delivered at the right time. The precision score was 92% accuracy for shipments arriving on time to the consumers. The recall score was 98%. The F1 scores was 99% for 0 and 95% for 1. The ethical issues are FedEx must make sure they respect their consumer's personal data and keep them satisfied by ensuring their packages are delivered to them on time. The model can be used for any business that has services and products.

References

Ahmed, I. (n.d.) FedEx Project One Dataset. Retrieved from [FedEx Project One | Kaggle](#) on November 10, 2022.

FedEx (n.d.) Company Structure and Facts. Retrieved from [Company Structure and Facts | FedEx](#), on November 10, 2022.

FedEx (n.d.) FedEx History. Retrieved from [Our History | FedEx](#), on November 10, 2022.

Appendix:

Table 1: Model Accuracy Scores

```
#Conduct the prediction tests and print the metrics.
y_pred = gb.predict(x_test)
print(confusion_matrix(y_test,y_pred))
print(accuracy_score(y_test,y_pred)*100)
print(classification_report(y_test,y_pred))
```

```
[[549291 11563]
 [ 3071 140509]]
97.92258749577675
```

	precision	recall	f1-score	support
0	0.99	0.98	0.99	560854
1	0.92	0.98	0.95	143580
accuracy			0.98	704434
macro avg	0.96	0.98	0.97	704434
weighted avg	0.98	0.98	0.98	704434