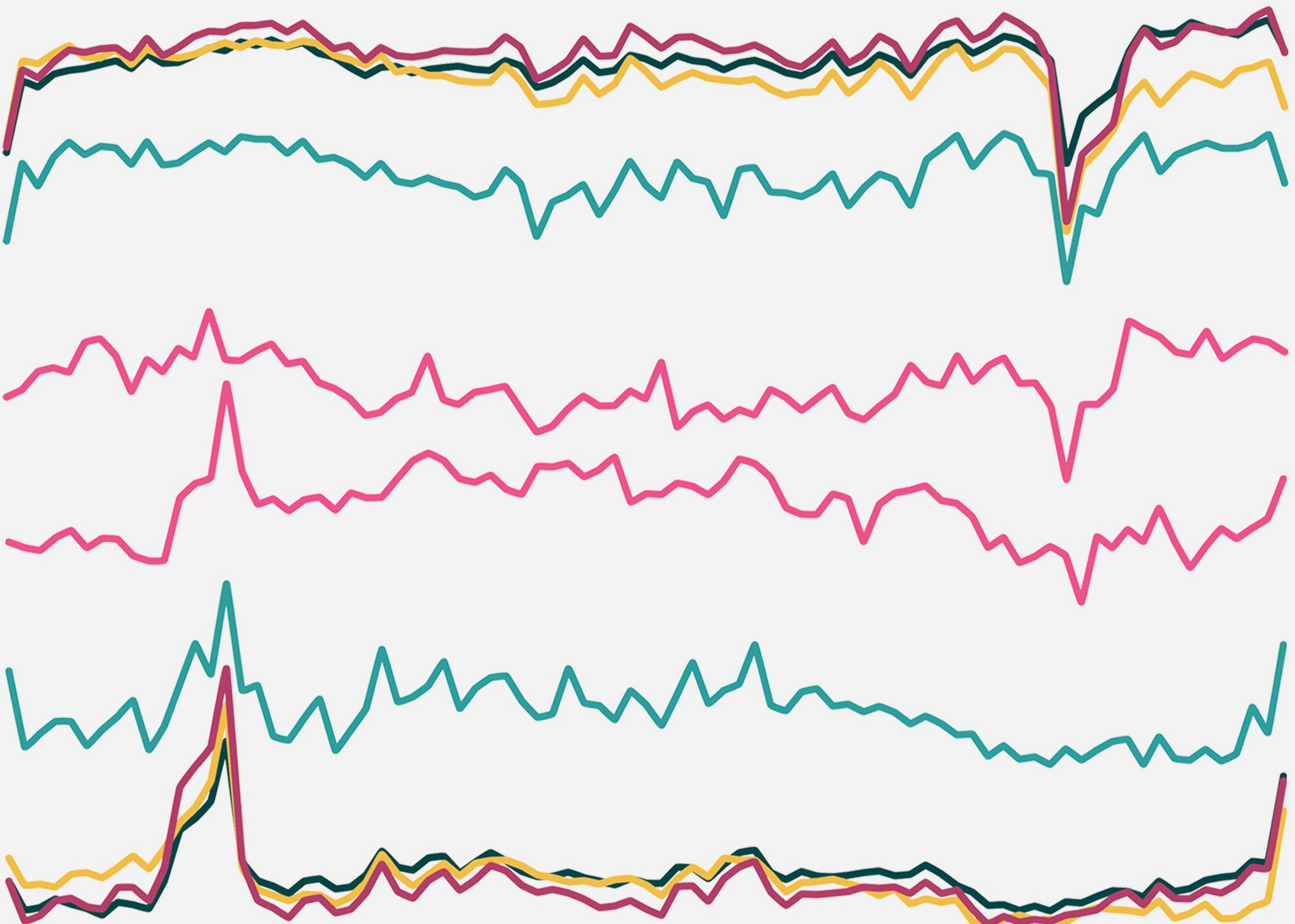


# Ali Naghibi

## Data Analyst



# Expert in Converting numbers into narratives.

Average Sold Product Value

\$403



Cumulative Discount

\$468,955



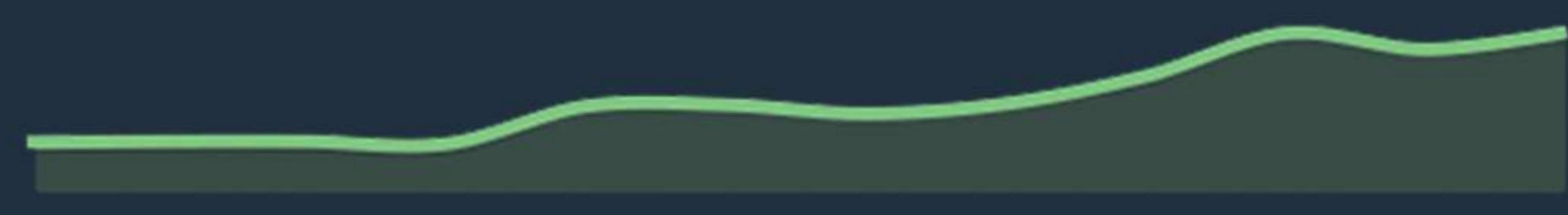
Unique Customers

18,553



Shipping Destinations

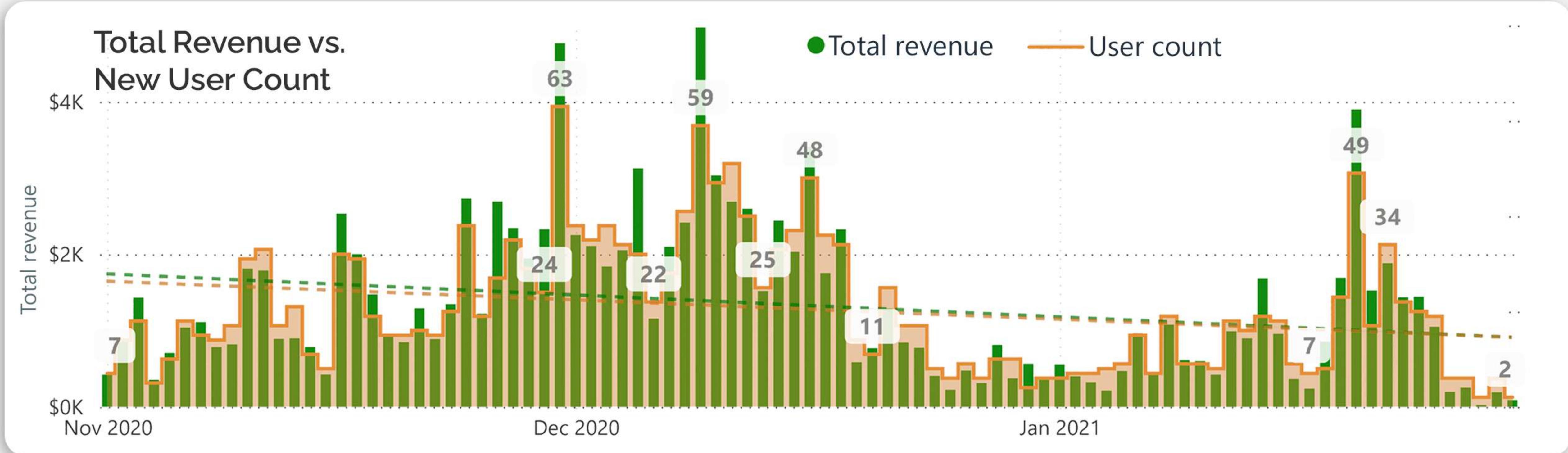
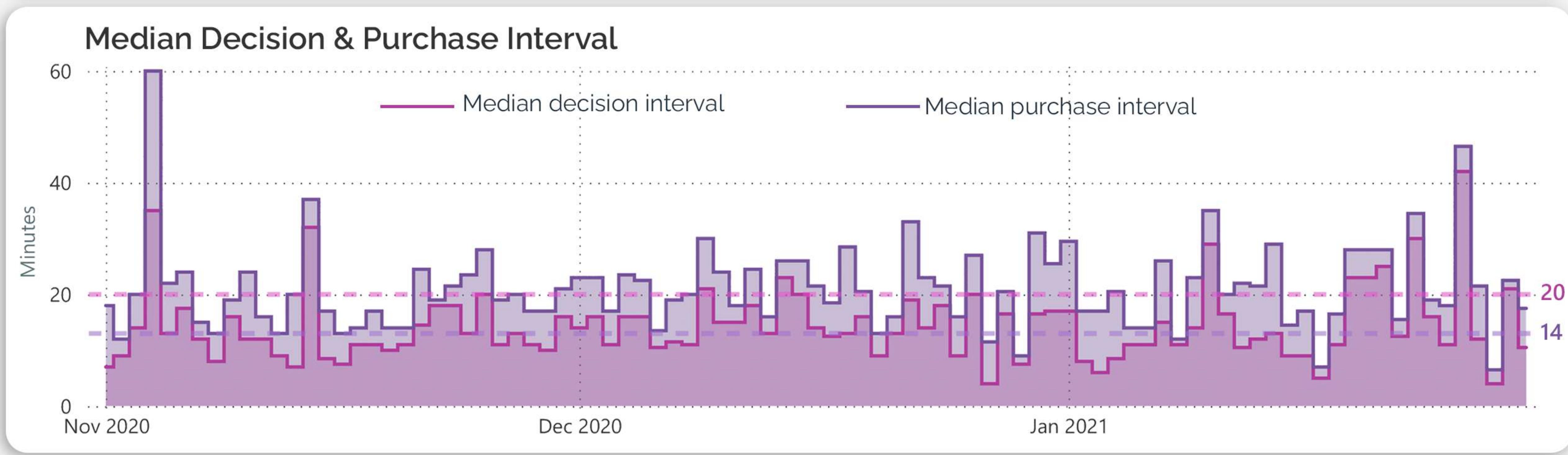
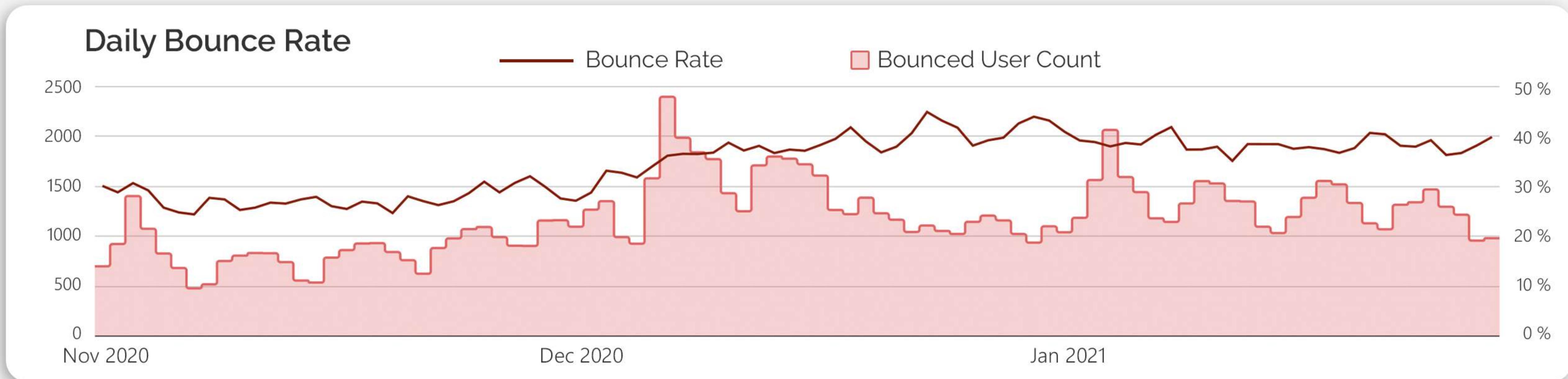
539



With expertise in tools such as PowerBI, Tableau, SQL, and Excel and a keen eye for detail, I turn raw data into actionable insights. Analytical by nature and creative at heart, I bridge the gap between data and action.

# Ad-hoc Analysis

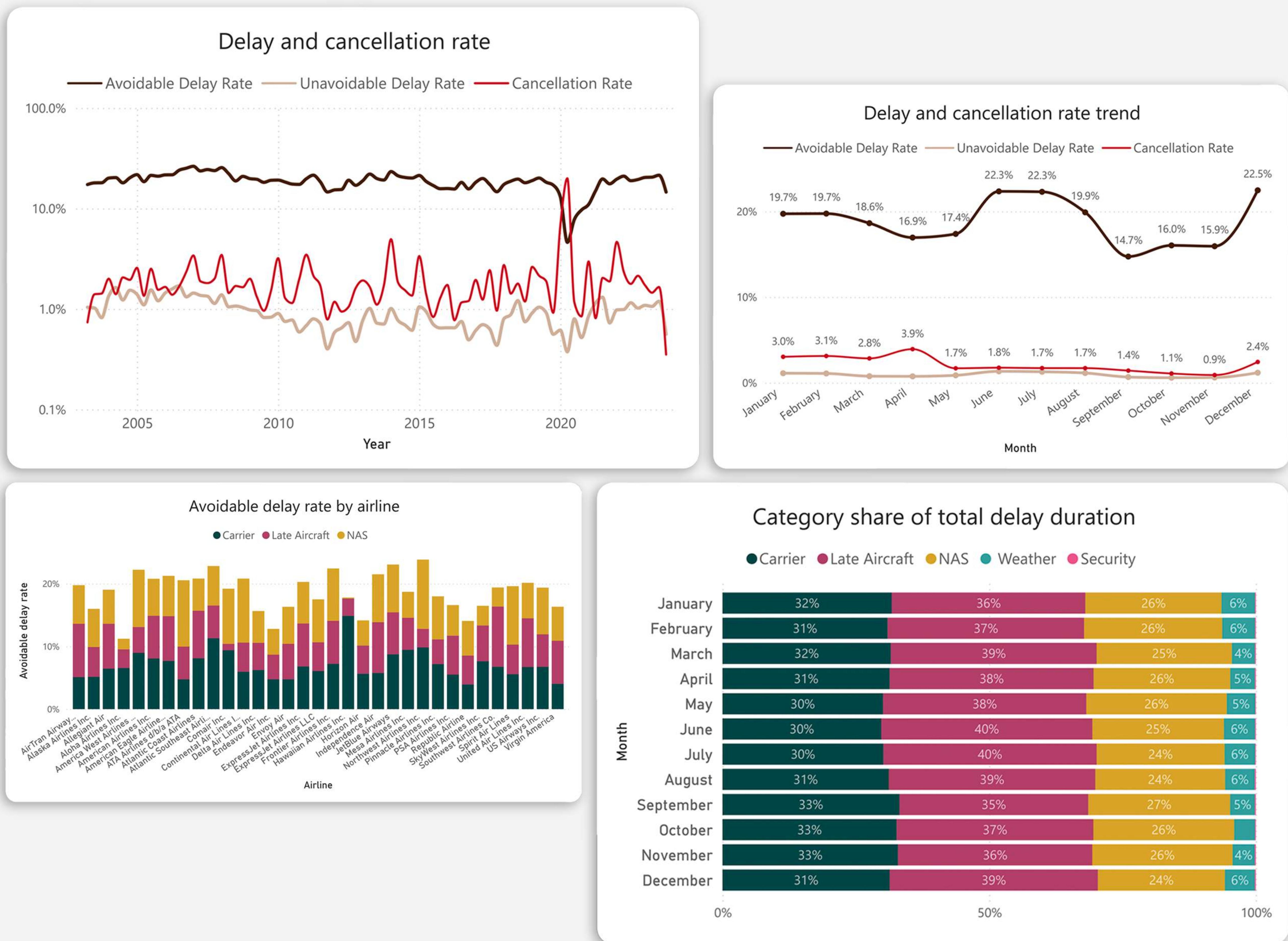
Exploration of key business drivers and providing clarity on customer actions, operational bottlenecks, and opportunities for value creation.



# Exploratory Data Analysis (EDA)

## United States Domestic Arrivals' Delay

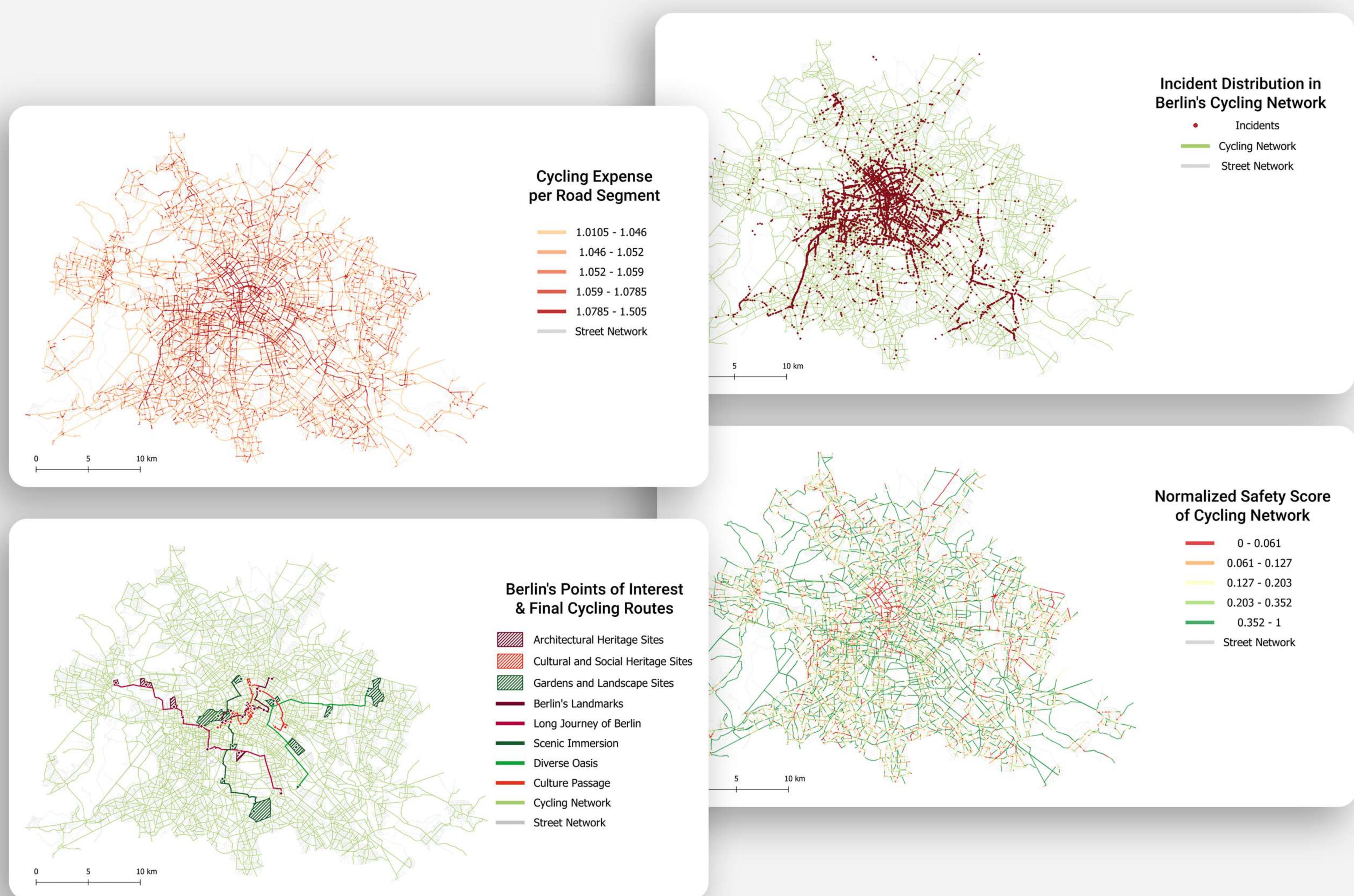
Comprehensive analysis of nationwide flight delay trends using large-scale datasets, highlighting key drivers, variability across airlines and seasons, and measurable impacts on efficiency.



# GIS Driven Safety Assessment and Route Modelling

## Developing Safe Cycling Routes for Exploring Berlin's Heritage Sites

Transforming crowd-sourced spatial datasets into reliable analytical framework for real-world mobility challenges.





# Dashboard Design

Showcasing the significance of numbers by bringing out perception and providing clarity with style and purpose.

# Documentation and Reporting

Clear, structured documentation outlining each stage of the analytical workflow, ensuring transparency, reproducibility, and effective communication with peers and stakeholders.

descriptive data including street and district name and when applicable, the street class in accordance with Berlin's street hierarchy.

3. **Heritage Site Dataset:** Obtained from Berlin State Heritage Office and curated from more than 12,000 entries on the original list. Each heritage site was georeferenced based on its street address and was appointed a category for the purpose of the research.

4. **Berlin Cycling Network:** Obtained from the Berlin Mobility Act (MobG BE) Transport, Climate Action, and Berlin Mobility Act (MobG BE) facilities in Berlin. This dataset main network for locating the and contains descriptive info neighborhood, and district.

All datasets were transformed to a coordinate system (WGS 84 / UTM Zone 33N, EPSG:25833) to enable spatial analysis as point layers into QGIS, while the hybrid dataset depending on the associated flow datasets were imported as line layers to them.

**Data Cleaning Procedures** included:

- Removal of duplicate lines in the dataset
- Filtering out mountain bike routes and ones in which the ride was either
- Giving a 10 meters buffer radius to factor in the width of bicycle lanes for further analysis.
- Joining the bike traffic data with the flow data and removing the overlapping lines
- Calculating the median for delay and appointing the amount to t

## Approach

The analysis is based on arrival flights of various carriers at U.S. public airports from June 2003 to December 2023. Data includes total incoming flights, delays, cancellations, diverted flights, reasons for each occurrence, and total delay duration. This data is combined with information on aviation facilities, such as location, state, and date of activation, for a comprehensive analysis. The data, sourced from the DOT's official website, has been cleaned and organised using BigQuery.

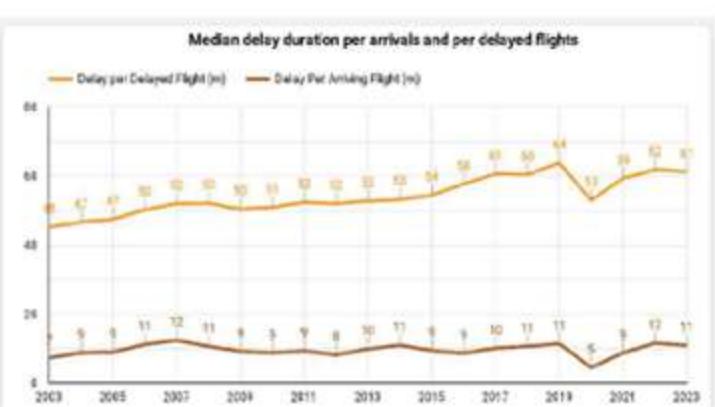
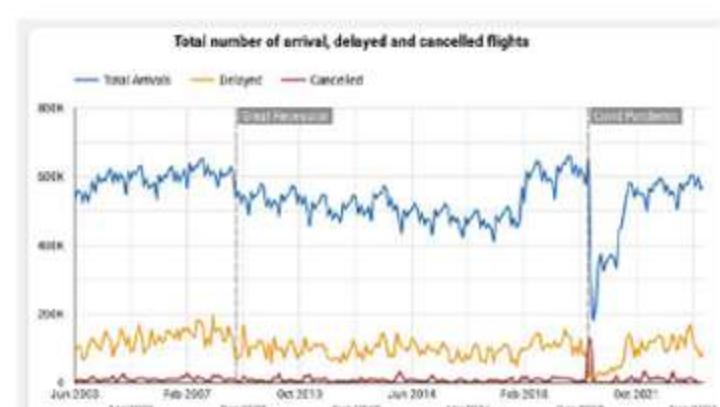
The analysis is divided into two major parts:

- **Avoidability:** Inspection based on avoidable and unavoidable roots of delays
- **Technicity:** Evaluation of performance of airlines, airports and air control (NAS) as functioning units of the system

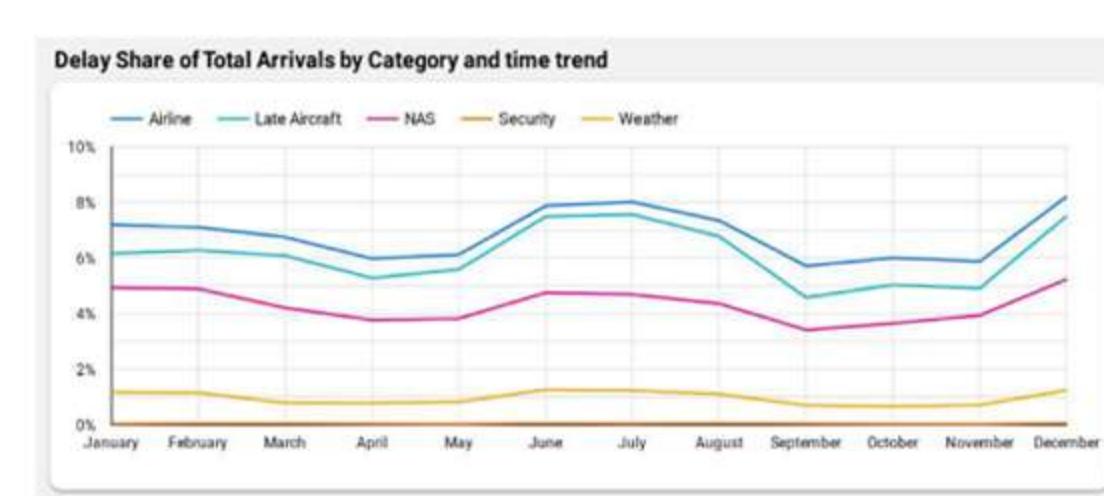
Metrics are converted into rates, shares, and percentages to account for the varying sizes of airlines and airports. Pareto charts are used to prioritise areas for improvement by weighing the effect of major elements in avoidable delays.

## Analysis

A historical overview shows a direct correlation between the number of flights and delays. Significant events such as the Great Recession and the COVID-19 pandemic have markedly affected flight numbers and delays. The median delay duration per total arrivals increased from 7 to 11 minutes, while the median delay duration per delayed arrival increased by 35%.



The breakdown of delayed arrivals indicates that the major causes of delay are late aircraft (6.62%), NAS (6.2%), and airlines (5.64%). Seasonal trends reveal a 5-6% rise in delays during summer and December, with a sharper increase in airline and late aircraft delays during these periods.



- Standardizing attribute naming conventions across datasets to allow relational joins in QGIS and SQL queries.

The integrated geodatabase allowed spatial joins between incidents, flow points, and heritage locations via a road network for route scoring and subsequent evaluation.

## 4.1.2 Development of the Safety Score

A **segment-based safety score** was developed for each segment in the Berlin cycling network, taking into account traffic volume, and road length, with adjustments for speed limits.

The base **Incident Density Score** was derived from the number of incidents per kilometer of the route.

## Incident Density Score

The incident density score is derived from the number of incidents per kilometer of the route. It is calculated by dividing the total number of incidents by the total length of the route. The higher the score, the more incidents occurred per kilometer. This score is used to identify segments of the route that are more prone to accidents. The score is then adjusted based on the number of cyclists using that segment. The final safety score is a weighted average of the segment-based scores and the incident density score.

These routes also include strategic rest points near cafés, public seating, and points of interest, ensuring that the journey is as much about the urban atmosphere as the heritage destinations themselves. Each individual route crosses the locations of another list and group in a few cases which reminiscences the user of their past journey on another route and gives them a complimentary sense of place and experience.

For each route, two alternatives were created and their Cost Score was compared to select the one with better outcome. The following table shows the results of this comparison and the more suitable routes as final outcomes.

Route	Length (km)	Avg. Cyclist Count	Avg. Incident density	Avg. Safety Score	Total Cost
Berlin's Landmarks 1	14.539	1135.5	0.027	0.260	120.57
Berlin's Landmarks 2	13.824	1008	0.028	0.296	<b>118.26</b>
Long Journey of Berlin 1	28.533	1015.9	0.018	0.326	202.842
Long Journey of Berlin 2	29.511	1124.8	0.017	0.387	<b>195.188</b>
Scenic Immersion 1	18.024	949.8	0.060	0.197	156.264
Scenic Immersion 2	19.613	880.9	0.077	0.172	<b>140.226</b>
Diverse Oasis 1	23.579	1151.4	0.015	0.383	<b>136.07</b>
Diverse Oasis 2	26.942	1094.9	0.018	0.346	179.744
Culture Passage 1	17.395	997.9	0.032	0.220	<b>160.352</b>
Culture Passage 2	21.999	1281	0.038	0.235	184.964

Table n: Evaluation of suggested routes based on safety and cost score, Author

As a result of this comparison, the following alternatives have been chosen as final solutions: **Berlin's Landmarks 2**, **Long Journey of Berlin 2**, **Scenic Immersion 2**, **Diverse Oasis 1** and **Culture Passage 1**.