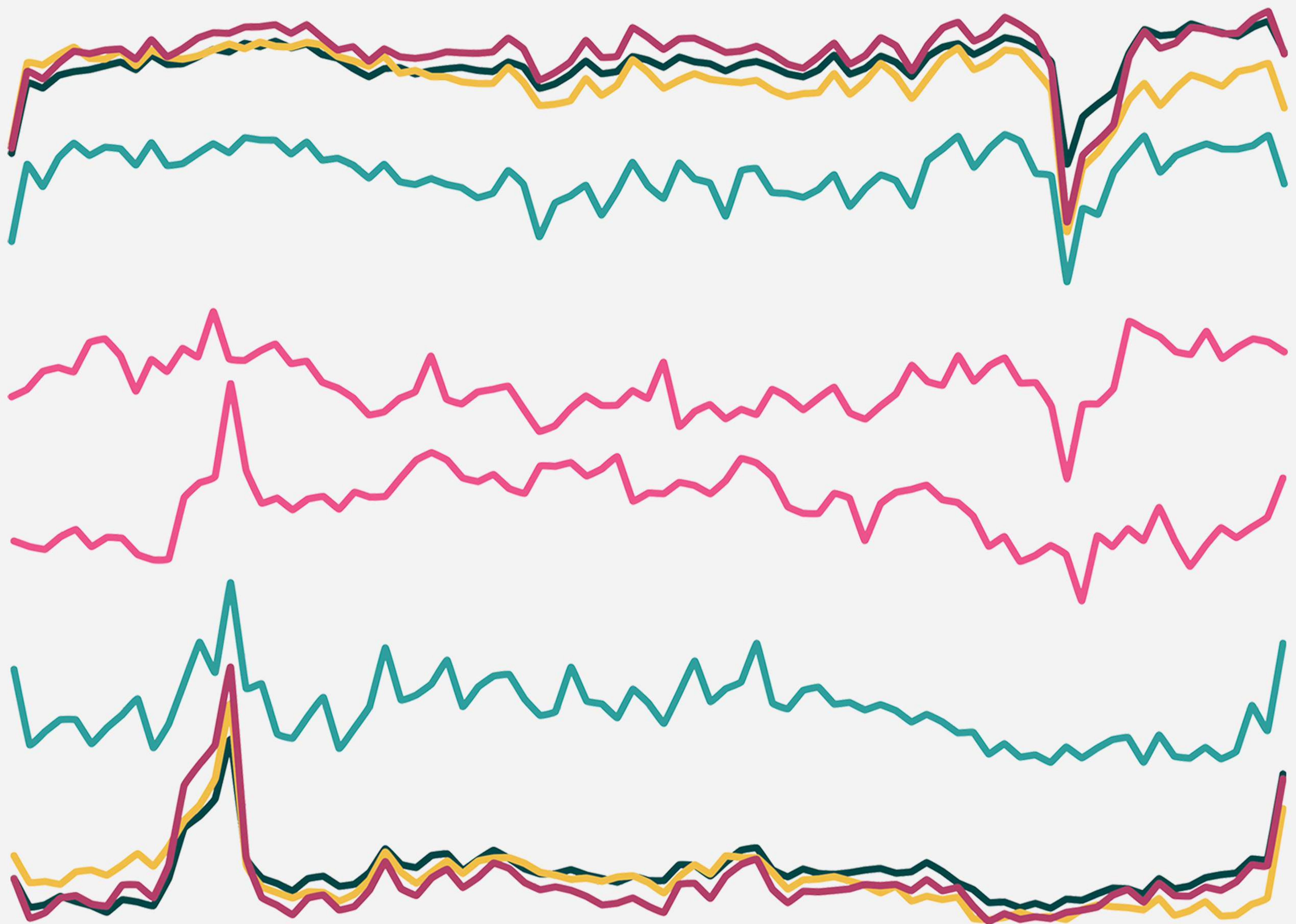


Ali Naghibi

Datenanalyst



Experte darin, Zahlen in Erzählungen zu verwandeln.

Average Sold Product Value

\$403



Cumulative Discount

\$468,955



Unique Customers

18,553



Shipping Destinations

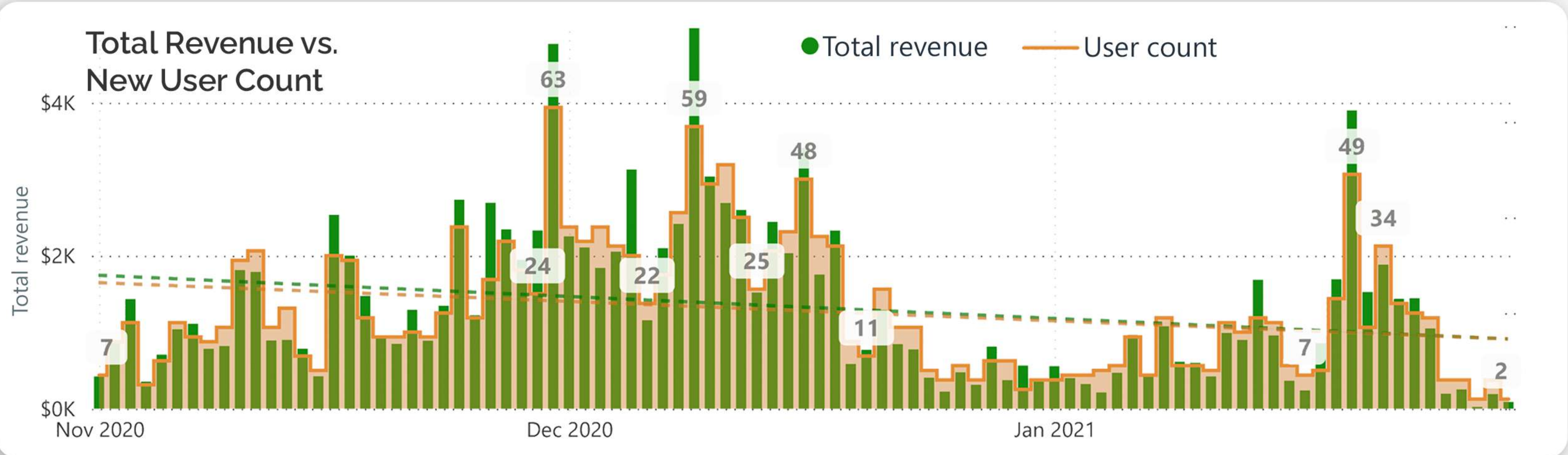
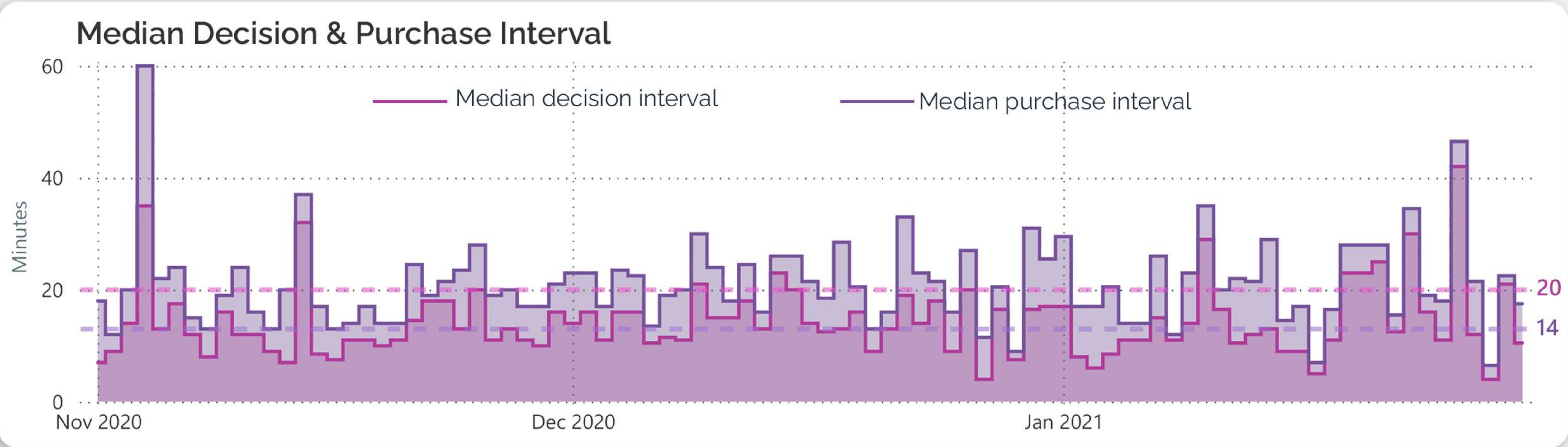
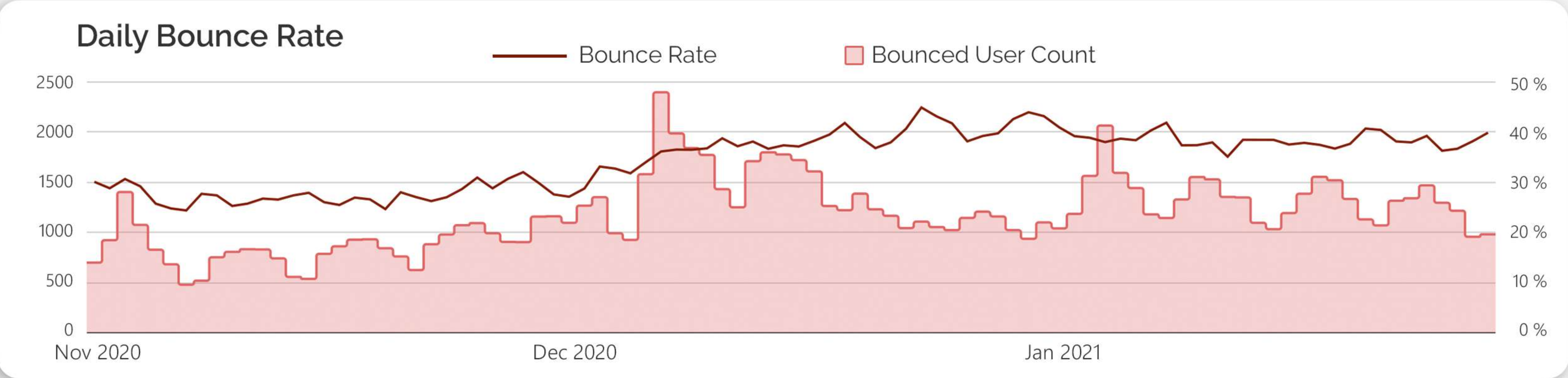
539



Mit fundierter Erfahrung in Power BI, Tableau, SQL und Excel sowie einem ausgeprägten Blick für Details, transformiere ich Rohdaten in umsetzbare Erkenntnisse. Analytisch im Kern und kreativ im Ansatz, überbrücke ich die Lücke zwischen Daten und konkreten Handlungsempfehlungen.

Ad-hoc-Analyse

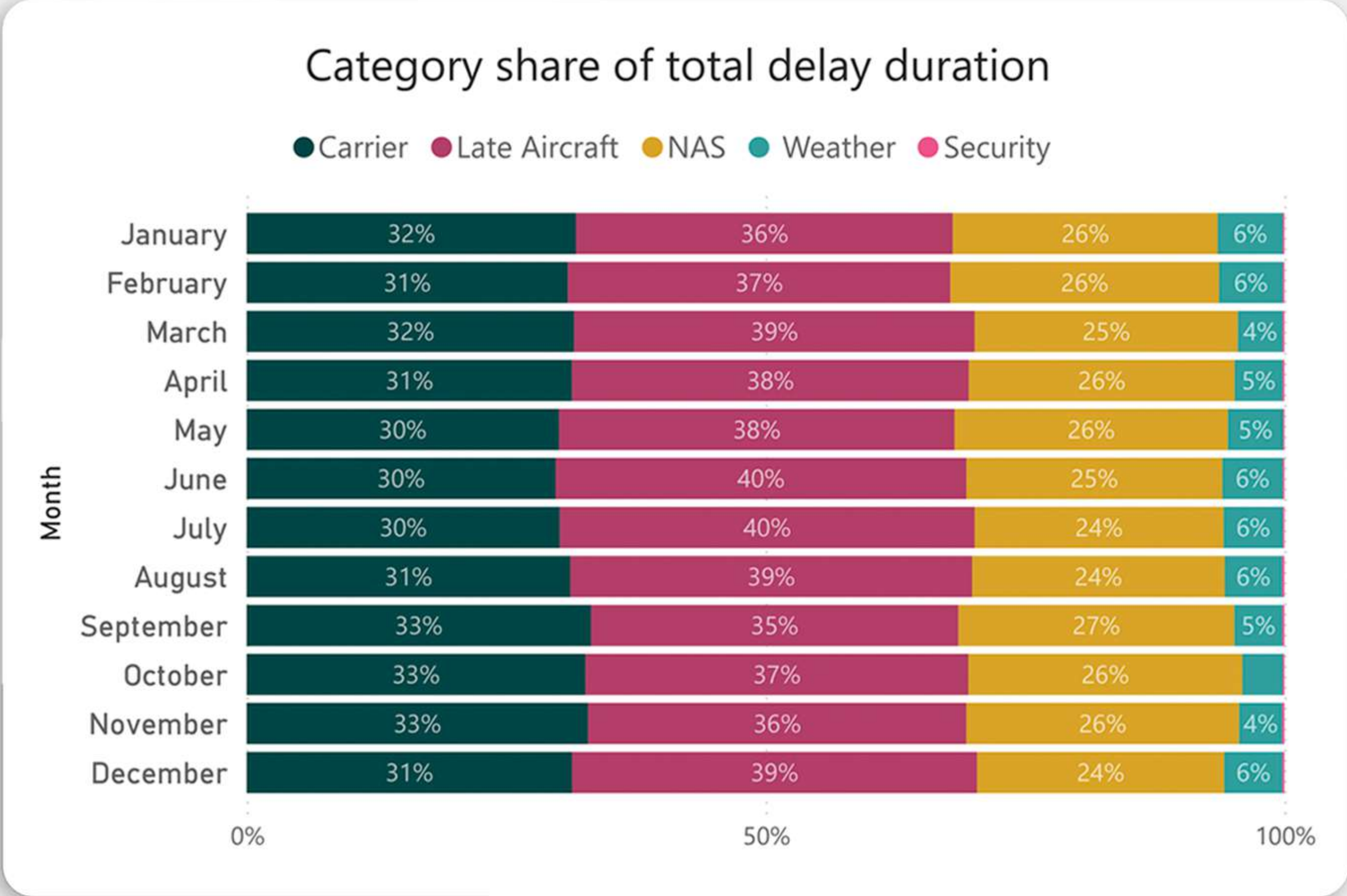
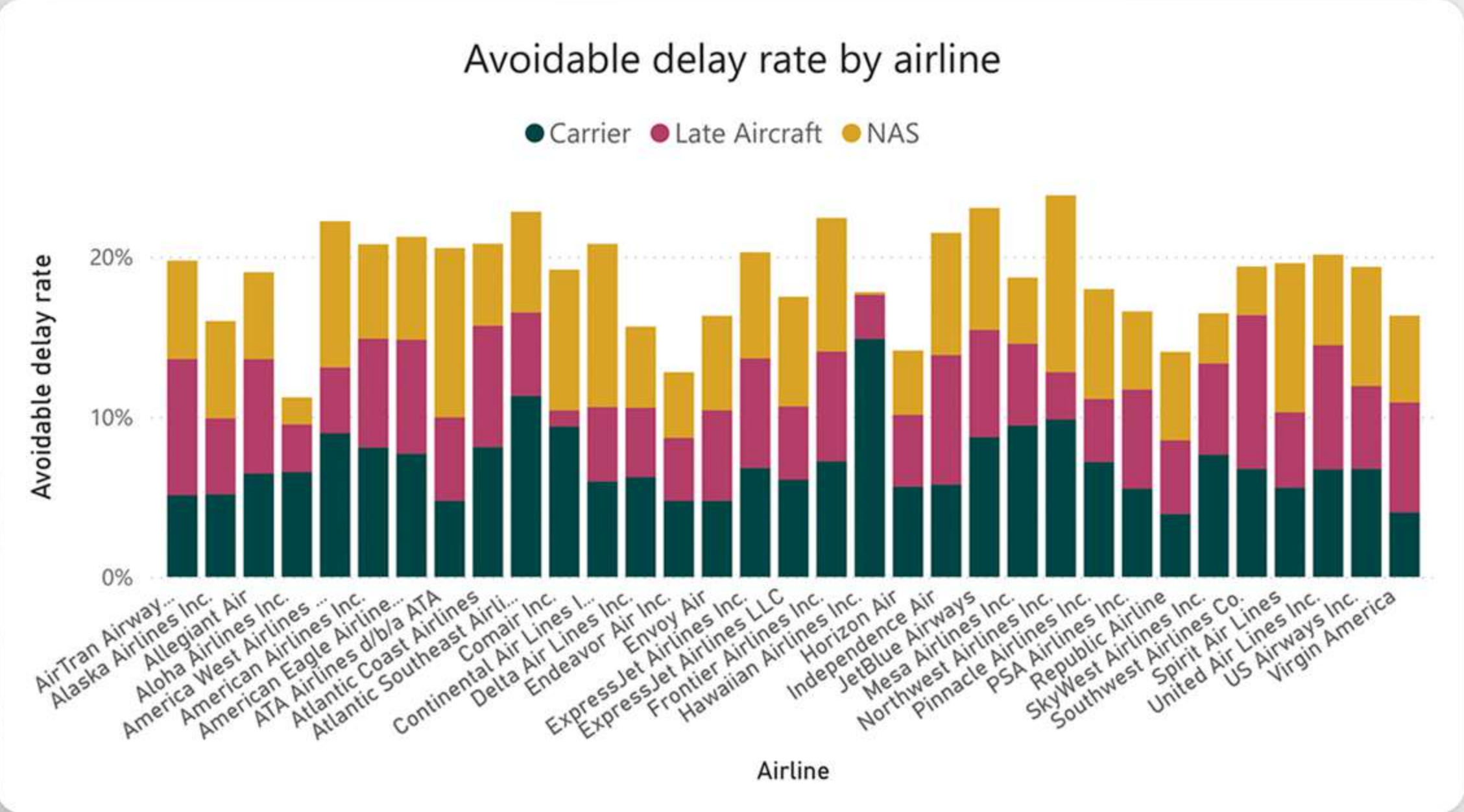
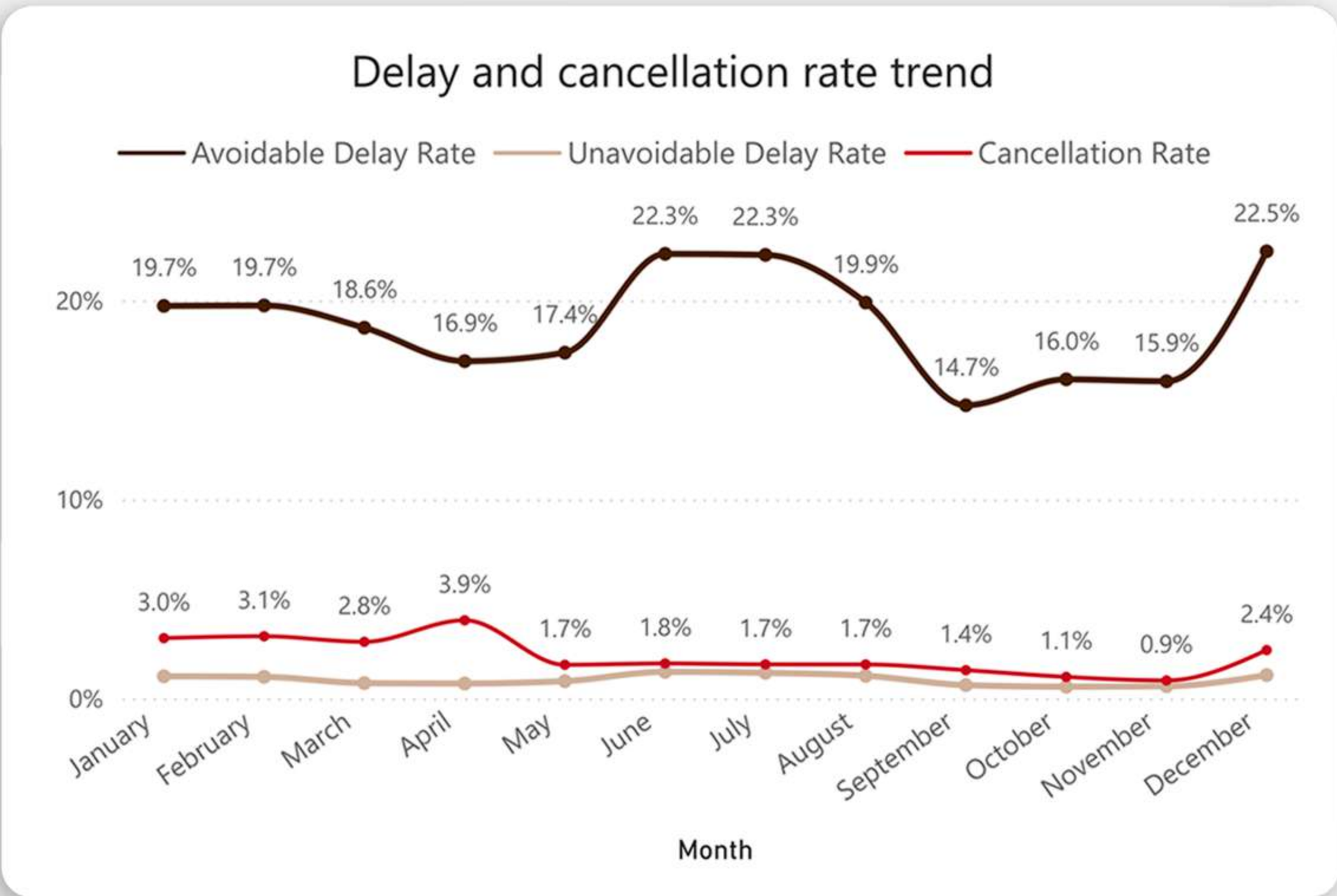
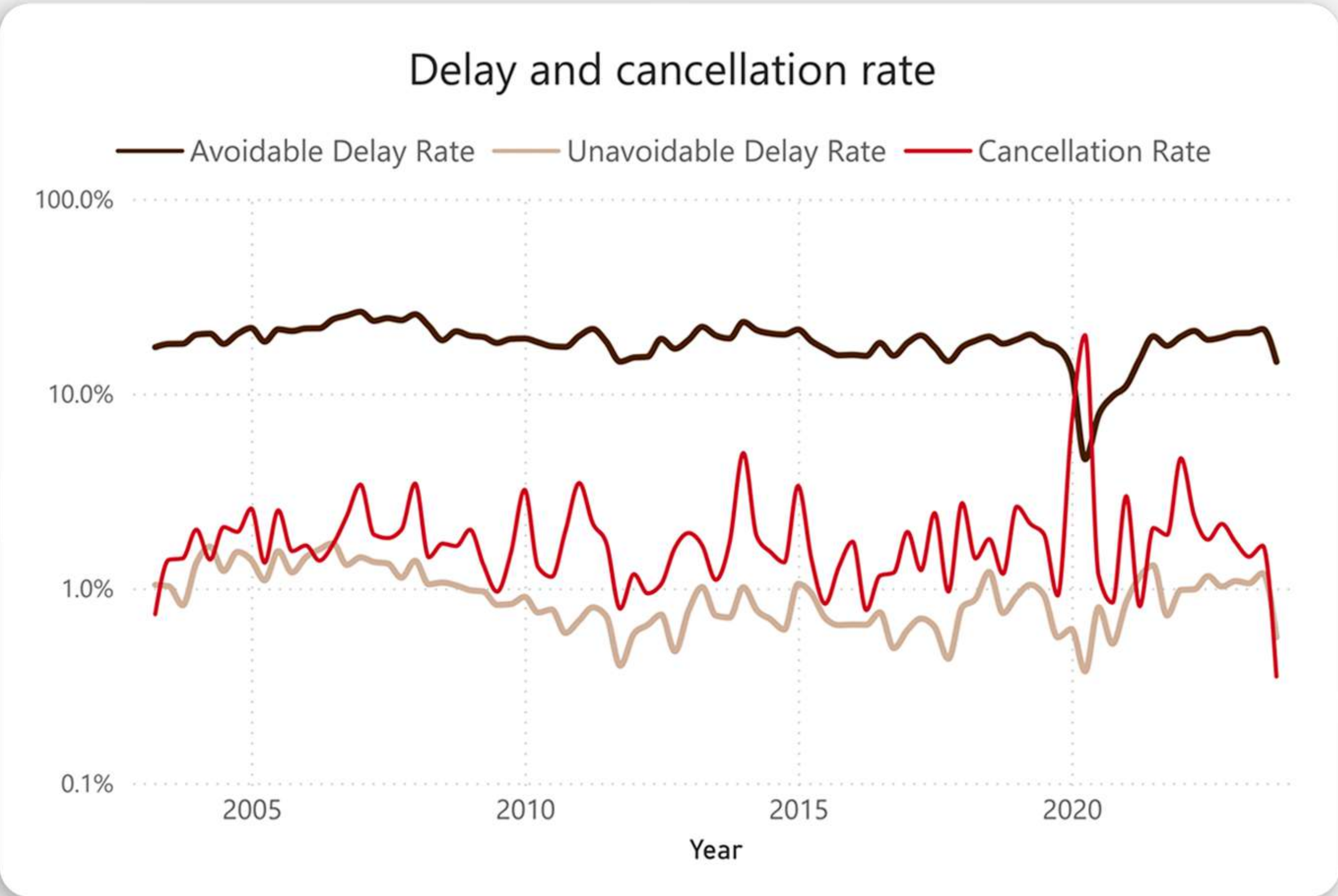
Auswertung von geschäftskritischen Faktoren und Untersuchung von Kundenverhalten, operativen Engpässen und Wertschöpfungspotenzialen.



Explorative Datenanalyse (EDA)

Verspätungen bei Inlandsankünften in den USA

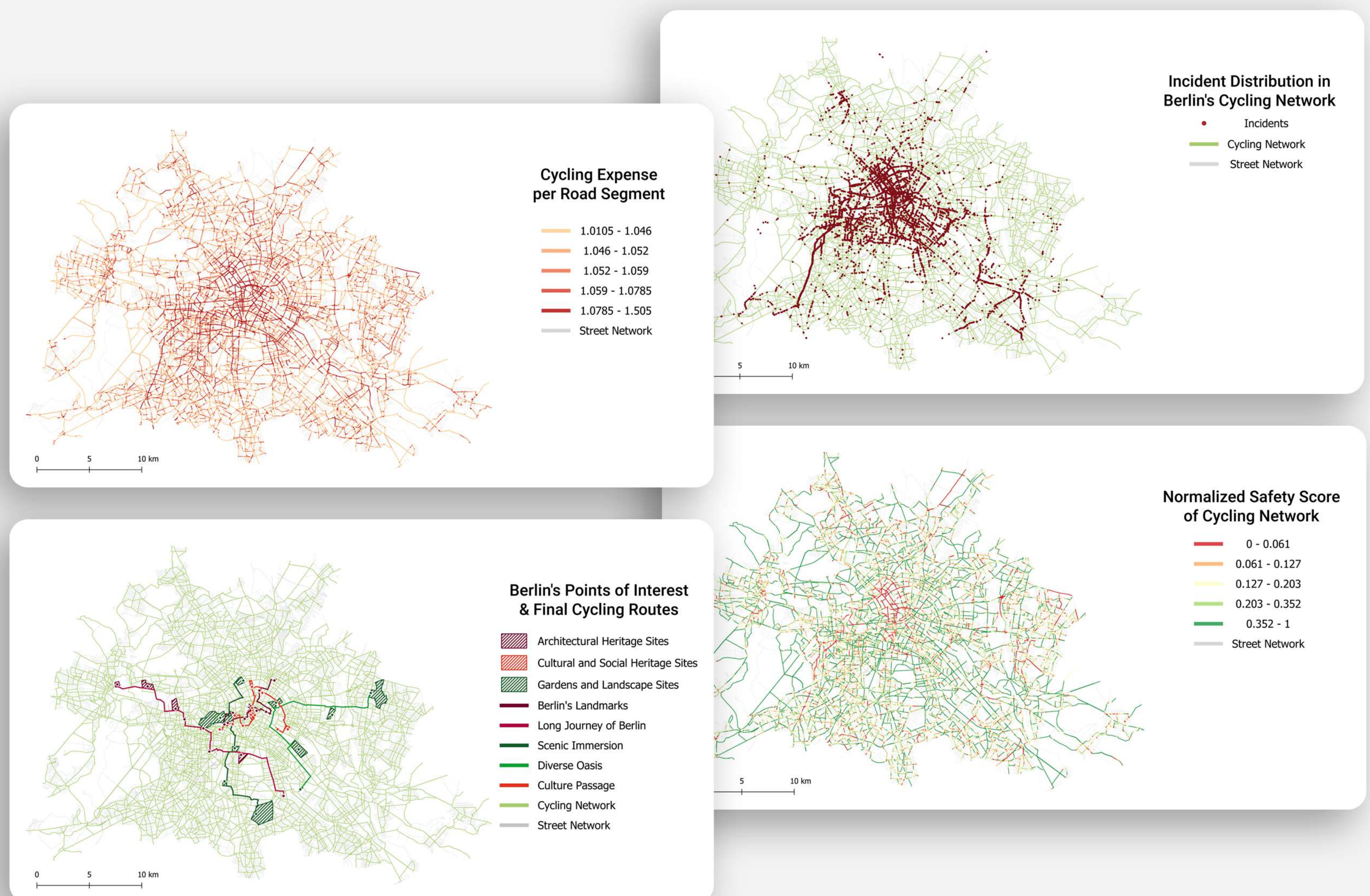
Umfassende Analyse landesweiter Flugverspätungstrends auf Basis umfangreicher Datensätze, mit Fokus auf zentrale Einflussfaktoren, Unterschiede zwischen Airlines und Jahreszeiten sowie messbare Auswirkungen auf die betriebliche Effizienz.



GIS-basierte Sicherheitsanalyse und Routenmodellierung

Entwicklung sicherer Fahrradrouten zur Erkundung von Berlins Kulturerbestätten

Umwandlung crowd-basierter Geodaten in ein belastbares Analyseframework zur Lösung realer Mobilitäts Herausforderungen.





Dashboard-Design

Veranschaulichung der Aussagekraft von Daten durch klare, zielgerichtete Visualisierungen, die Erkenntnisse sichtbar und verständlich machen.

Dokumentation und Reporting

Klare und strukturierte Dokumentation des Workflows, die Transparenz, Reproduzierbarkeit und eine effiziente Kommunikation mit Teammitgliedern und Stakeholdern sicherstellt.

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 State Office for Transport, Climate Action, and the Berlin Mobility Act (MobG BE) to include all cycling routes and facilities in Berlin. This dataset was used as the main network for locating the routes and contains descriptive information on the neighborhood, and district.

All datasets were transformed to a coordinate system (Zone 33N, EPSG:25833) to enable spatial analysis. The point layers as point layers into QGIS, while the flow datasets were imported as line layers to them.

Data Cleaning Procedures included:

- Removal of duplicate lines in the dataset
- Filtering out mountain bikes and e-bikes from the dataset
- Giving a 10 meters buffer radius around the points to factor in the width of bicycle lanes
- Joining the bike traffic data with the street data and removing the overlapping lines
- Calculating the median for daily flows and appointing the amount to the street

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
- **Technicality:** 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%.

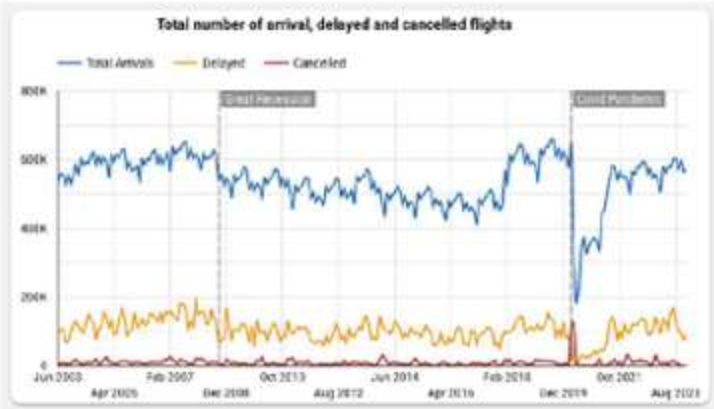


Fig. 1: Total number of arrivals, delays and cancellations from 2003-2023

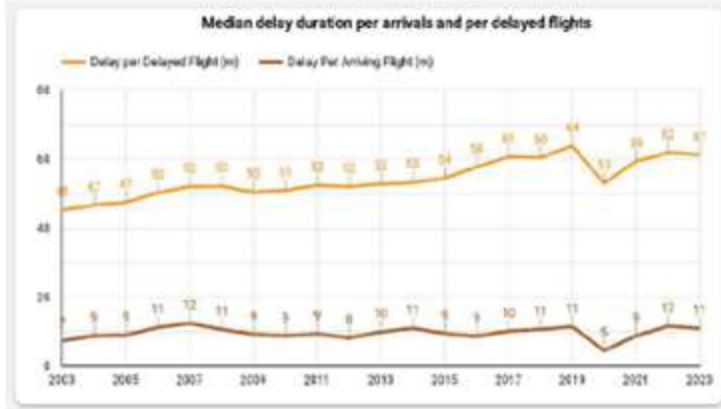


Fig. 2: Median delay duration by arrival and delayed arrival

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.

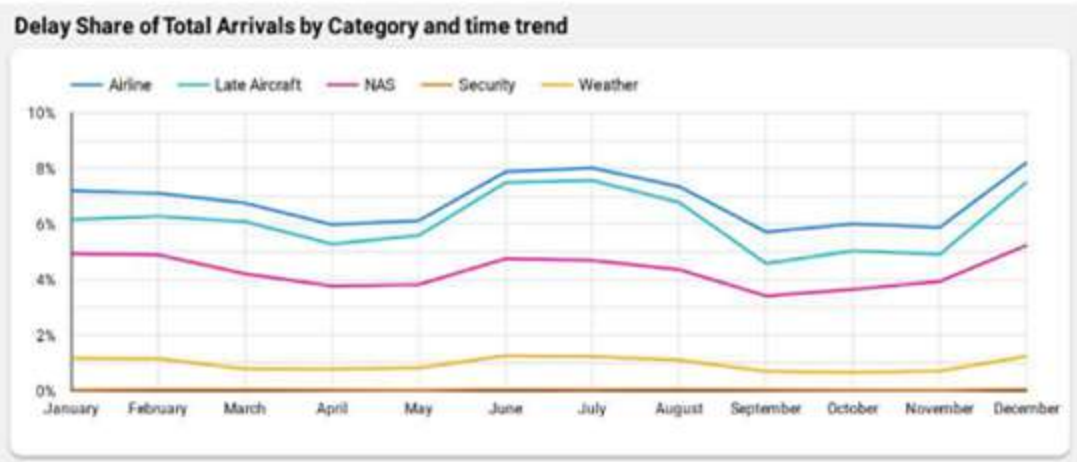


Fig3: Seasonal trend of delays by reason

- 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. This enabled the calculation of route scores for route scoring and subsequent evaluation.

4.1.2 Development of the Safety Score

A **segment-based safety score** was calculated for each segment in the Berlin cycling network. The score was based on incident volume, and road length, with adjustments for road type and traffic volume.

The base **Incident Density Score** was calculated as follows:

Incident Density Score

derived from

derived from

ant the

y incident

counted

) to prevent

investment

kely to

ferences

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	118.26
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	195.188
Scenic Immersion 1	18.024	949.8	0.060	0.197	156.264
Scenic Immersion 2	19.613	880.9	0.077	0.172	140.226
Diverse Oasis 1	23.579	1151.4	0.015	0.383	136.07
Diverse Oasis 2	26.942	1094.9	0.018	0.346	179.744
Culture Passage 1	17.395	997.9	0.032	0.220	160.352
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**.