



A.D. 1308
unipg
DIPARTIMENTO
DI INGEGNERIA

Tesina Finale di
Models And Algorithms for Data Visualization
Corso di Laurea Magistrale in Ingegneria Informatica e Robotica – A.A. 2024-2025
DIPARTIMENTO DI INGEGNERIA

Docente
Prof. Giuseppe LIOTTA

**Sistema di Visualizzazione per i Dati sul Traffico Aereo
Internazionale USA-Estero**



studenti

374435	Martin	Arsoski	martin.arsoski@studenti.unipg.it
374500	Michele	Mariotti	michele.mariotti@studenti.unipg.it
375564	Riccardo	Nicolini	riccardo.nicolini1@studenti.unipg.it

22 febbraio 2025

0. Indice

1	Introduzione	2
2	Raccolta dei Requisiti	4
2.1	Utenti Finali	4
2.2	Data Modeling	5
2.2.1	Data extraction	5
2.2.2	Data cleaning and transformation	8
2.2.3	Definizioni dei tipi di Dataset e degli Attributi	11
2.3	Task Modeling	12
3	Progettazione	14
3.1	Progettazione della Visualizzazione	14
3.1.1	Choropleth Map	14
3.1.2	Node-Link Connection Map	15
3.2	Progettazione dell'interazione	16
3.3	Scelte Tecnologiche ed Architetture	16
4	Realizzazione	18

1. Introduzione

Il progetto è finalizzato allo sviluppo di un sistema interattivo per l'analisi del traffico aereo internazionale tra gli Stati Uniti e il resto del mondo, con l'obiettivo di rendere i dati facilmente accessibili e utili a diverse categorie di utenti.

L'obiettivo principale è fornire agli utenti uno strumento che permetta di:

- consultare la connettività tra gli Stati Uniti e i diversi Stati esteri;
- esplorare i dati relativi al traffico aereo, con particolare attenzione ai flussi di passeggeri e voli suddivisi per anno e mese;
- analizzare i collegamenti aerei con il maggiore numero di traffico.

Il dataset utilizzato per il progetto è "U.S. International Air Traffic Data (1990-2020)", disponibile sulla piattaforma Kaggle.

Il sistema di visualizzazione sviluppato consente di osservare i dati, suddivisi per anno e mese, attraverso due idiomi visuali: una choropleth map, utilizzata per visualizzare il numero di collegamenti esteri di ciascuno Stato US, e una node-link connection map, utilizzata per rappresentare i collegamenti tra gli Stati US e quelli esteri, includendo anche il numero di voli e passeggeri associati a ciascun collegamento.

La visualizzazione di questi dati permette di ottenere una comprensione immediata e chiara delle connessioni internazionali, superando le limitazioni delle analisi tradizionali statiche. Inoltre, si rivela uno strumento utile per supportare decisioni strategiche nel settore aereo, offrendo un valore concreto ad analisti, autorità governative e gestori aeroportuali.

Le sezioni seguenti sono organizzate come segue:

- **Raccolta dei Requisiti:** La sezione si apre con l'identificazione degli utenti finali e delle loro necessità principali. Successivamente, viene descritto il processo di modellazione dei dati, che include le attività di estrazione, pulizia e trasformazione, e vengono definite le tipologie di dataset e gli attributi utilizzati. Infine, viene illustrato il task modeling, che specifica le operazioni analitiche che il sistema deve supportare, come l'individuazione dei collegamenti aerei e la ricerca delle connessioni più trafficate.
- **Progettazione:** In questa sezione vengono descritte le decisioni progettuali che hanno portato alla definizione delle visualizzazioni e delle interazioni. Si approfondiscono gli idiomi visuali scelti e vengono delineati gli strumenti di interazione, insieme alle scelte tecnologiche e architetturali.
- **Realizzazione:** La parte conclusiva illustra i dettagli tecnici dello sviluppo, mettendo in evidenza come il sistema permetta agli utenti di navigare tra i dati, selezionare stati e periodi temporali, e ottenere informazioni visive chiare e dinamiche.

2. Raccolta dei Requisiti

Il dataset utilizzato per questo progetto è stato ottenuto da Kaggle, una piattaforma di condivisione e analisi di dati. In particolare, il modello preso in considerazione è "*U.S. International Air Traffic Data (1990-2020)*".

Questo dataset fornisce informazioni dettagliate sul traffico aereo internazionale relativo ai voli diretti tra gli aeroporti degli Stati Uniti e quelli esteri, dal 1990 al 2020.

2.1 Utenti Finali

Il sistema di visualizzazione è progettato per un'ampia gamma di utenti finali interessati all'analisi e al monitoraggio del traffico aereo internazionale tra gli Stati Uniti e il resto del mondo.

Chi sono?

I principali utenti finali del sistema includono:

- **Analisti del settore aereo:** Professionisti incaricati di analizzare il traffico aereo per la pianificazione strategica e operativa.
- **Autorità governative:** Enti come il Dipartimento dei Trasporti degli Stati Uniti (DOT) e la Federal Aviation Administration (FAA), responsabili dell'elaborazione di regolamenti e politiche di trasporto aereo.
- **Gestori di aeroporti:** Responsabili della gestione operativa degli aeroporti e della pianificazione delle rotte.
- **Ricercatori e accademici:** Studiosi nei campi del trasporto aereo, della logistica e dell'economia dei trasporti.

- **Agenzie di turismo e marketing:** Professionisti del settore turistico che utilizzano i dati per identificare mercati potenziali e pianificare campagne di marketing mirate.

Quali sono le loro necessità?

Gli utenti finali hanno esigenze specifiche che il sistema di visualizzazione deve soddisfare per supportare le loro attività professionali:

- **Esplorazione dei dati:** Accesso rapido ai dati sul traffico aereo in base a criteri temporali e geografici specifici.
- **Analisi dei collegamenti e della connettività:** Individuazione dei collegamenti più trafficati e delle opportunità di espansione dei collegamenti.
- **Monitoraggio delle prestazioni:** Tracciamento delle tendenze storiche dei volumi di passeggeri e voli.
- **Analisi comparativa:** Confronto tra volumi di traffico di diversi Stati e periodi temporali

2.2 Data Modeling

2.2.1 Data extraction

Il dataset è strutturato in due tabelle principali (file .csv) che rappresentano due aspetti del traffico aereo: "**Departures**" e "**Passengers**". Ogni tabella contiene dati aggregati per mese e anno, suddivisi per tratta, compagnia aerea e tipo di volo (programmato o charter). Entrambi i dataset descrivono i collegamenti aerei tra gli Stati Uniti e destinazioni internazionali, ma si concentrano su aspetti differenti:

Il dataset "**Departures**" include il numero di voli effettuati da una determinata compagnia aerea tra un aeroporto gateway negli Stati Uniti e un aeroporto gateway internazionale.

Il dataset "**Passengers**" fornisce il numero di passeggeri trasportati da una compagnia aerea su voli tra un aeroporto negli Stati Uniti e un aeroporto internazionale.

Gli attributi del dataset possono essere interpretati come segue nella tabella 2.1.

Attributo	Descrizione
data_dte	Primo giorno del mese in cui i dati sono stati registrati.
Year	Anno del mese di riferimento dei dati.
Month	Mese del periodo di riferimento dei dati.
usg_apt_id	Identificativo univoco dell'aeroporto gateway negli Stati Uniti.
usg_apt	Codice dell'aeroporto gateway negli Stati Uniti (IATA).
usg_wac	Codice dell'area geografica mondiale per gli Stati Uniti.
fg_apt_id	Identificativo univoco dell'aeroporto gateway internazionale.
fg_apt	Codice dell'aeroporto gateway internazionale (IATA).
fg_wac	Codice dell'area geografica mondiale per l'estero.
airlineid	Identificativo univoco della compagnia aerea.
carrier	Codice della compagnia aerea (IATA).
carriergroup	Gruppo della compagnia aerea (1 per domestica USA, 0 per straniera).
type	Identifica il dataset ("Passengers" o "Departures").
Scheduled	Numero di passeggeri o voli programmati.
Charter	Numero di passeggeri o voli charter.
Total	Numero totale di passeggeri o voli.

Tabella 2.1: Descrizione degli attributi dei dataset "Departures" e "Passengers"

```
05/01/2006,2006,5,12016,GUM,5,13162,MAJ,844,20177,PFQ,1,Departures,0,10,10
05/01/2003,2003,5,10299,ANC,1,13856,OKO,736,20007,5Y,1,Departures,0,15,15
03/01/2007,2007,3,10721,BOS,13,12651,KEF,439,20402,GL,1,Departures,0,1,1
12/01/2004,2004,12,11259,DAL,74,16271,YYZ,936,20201,AMQ,1,Departures,0,1,1
05/01/2009,2009,5,13303,MIA,33,11075,CMW,219,21323,5L,0,Departures,0,20,20
```

Figura 2.1: Esempio tuple dataset "Departures"

Per migliorare la comprensibilità e l'analisi del traffico aereo, è stato utilizzato un ulteriore file chiamato **iata_codes.csv** (anch'esso ottenuto da Kaggle). Questo dataset contiene informazioni geografiche sugli aeroporti, organizzate nel seguente modo:

Attributo	Descrizione
iata_code	Codice univoco che identifica un aeroporto nel mondo.
Country_Name	Paese di appartenenza (o stato US nel caso degli Stati Uniti).
Country_Code	codice ISO-2 del paese (o stato US nel caso degli Stati Uniti).
Continent	Continente in cui si trova l'aeroporto.

Tabella 2.2: Descrizione degli attributi del dataset "iata_codes"

Queste informazioni sono state integrate nei dataset principali "**Departures**" e "**Passengers**" per arricchire i dati con il nome dello stato e del continente relativo a ciascun aeroporto. Grazie a questa integrazione, ogni tratta aerea nel dataset finale è associata con informazioni geografiche complete, che permettono di analizzare il traffico aereo in relazione agli stati e ai continenti, piuttosto che tra gli aeroporti.

```
LIRV,Italy,IT,Europe  
PEG,Italy,IT,Europe  
LJCE,Slovenia,SI,Europe  
LJU,Slovenia,SI,Europe  
MBX,Slovenia,SI,Europe  
POW,Slovenia,SI,Europe  
LJSG,Slovenia,SI,Europe  
LKCS,Czech Republic,CZ,Europe
```

Figura 2.2: Esempio tuple dataset "iata_codes"

2.2.2 Data cleaning and transformation

La pulizia e la trasformazione dei dati sono fasi cruciali per garantire che i dati siano pronti per un'analisi accurata e una visualizzazione significativa. Nel caso in esame, è stato effettuato un lavoro sui dataset per ottenere una rappresentazione utile allo scopo del progetto e di facile comprensione.

Data cleaning

La pulizia dei dati è il primo passaggio per garantire che i dati siano privi di errori, valori mancanti o formati incoerenti. Sono stati eseguiti i seguenti passaggi:

1. **Rimozione dei duplicati:** la prima verifica riguarda la presenza di duplicati all'interno dei dataset `departures` e `passengers`. Data l'assenza di duplicati non effettuiamo modifiche.
2. **Modifica colonne non utili o ridondanti:** alcuni attributi nei due dataset sono ridondanti, come per esempio (`data_dte`) e la coppia (`Year`, `Month`).

Altre colonne sono state rimosse, poiché non risultano utili ai fini dell'analisi visuale prevista, come per esempio le colonne (`charter`) e (`Total`), dal momento che l'interesse principale è focalizzato esclusivamente sui voli programmati (non `charter`). I voli `scheduled` rappresentano la componente stabile e regolare del traffico aereo commerciale, in quanto seguono orari prevedibili e rispondono a una domanda costante di trasporto. Al contrario, i voli `charter`, essendo operazioni occasionali, non riflettono le principali dinamiche del traffico e potrebbero introdurre rumore nei dati, rendendo meno chiara l'analisi.

Si precisa che lo Stato del Delaware non dispone di voli programmati ma solo di voli `charter`. Pertanto, il numero totale di stati statunitensi considerati nell'analisi è pari a 49.

Data transformation

Una volta puliti i dati, è stato necessario trasformarli per adattarli al modello desiderato, eseguendo `join`, raggruppamenti e filtri.

1. **Join delle tabelle `Departures` e `Passengers`,** Un primo `join` è stato effettuato tra le tabelle `Departures` e `Passengers`, creando un dataset unico in cui, per ciascuna tratta, venivano associati il numero di voli e il numero di passeggeri. Il tipo di `join` utilizzato è stato un `inner join`, il che significa che

solo le righe con corrispondenze in entrambe le tabelle sono state mantenute, escludendo quindi i campioni privi di dati in uno dei due dataset.

2. **Raggruppamento per compagnie:** I dati sono stati raggruppati per compagnia aerea, sommando il numero di passeggeri e voli effettuati da ciascuna compagnia su ogni tratta, mese e anno. In questo modo, tutte le informazioni relative alle operazioni di diverse compagnie aeree su una specifica rotta sono state aggregate in un'unica tupla. La scelta di raggruppare i dati per tratta, indipendentemente dalle compagnie aeree, permette di concentrarsi sul traffico complessivo di passeggeri e voli, fornendo una misura più diretta della domanda e della capacità delle rotte.
3. **Join con il dataset `iata_codes.csv`:** Poiché gli aeroporti sono identificati tramite i codici IATA, è stato necessario associare ciascun aeroporto al relativo stato e continente. A tal fine, è stato utilizzato il dataset `iata_codes`, che contiene il nome dello stato, il codice ISO e il continente per ogni aeroporto. Attraverso un'operazione di merge, queste informazioni sono state integrate nei dati, sia per gli aeroporti situati negli Stati Uniti che per quelli internazionali.
4. **Raggruppamento per Stati:** Per semplificare l'analisi e ottenere una visione più significativa dei flussi aerei, i dati sono stati aggregati per stato di partenza e destinazione, insieme a anno e mese. In questo modo, invece di analizzare singoli aeroporti, il dataset riflette il traffico aereo tra i diversi stati, con il numero di voli e passeggeri sommati per ciascun gruppo. La scelta migliora anche la chiarezza visiva, consentendo di rappresentare in modo più immediato le principali connessioni geografiche e i flussi di traffico, rendendo l'analisi più comprensibile ed efficace.

Alla fine, è stato generato un file JSON suddiviso in due insiemi principali: l'insieme **states** e l'insieme **links**.

Nell'insieme **states** sono presenti oggetti che rappresentano gli Stati coinvolti nel traffico aereo. Gli attributi di ciascun oggetto includono il nome dello Stato, il codice ISO-2 e il continente di appartenenza.

Nell'insieme **links** sono invece presenti oggetti che rappresentano le tratte aeree tra gli Stati. Ogni oggetto contiene gli attributi relativi all'anno e al mese di riferimento, allo Stato US e lo Stato straniero (identificati dai rispettivi codici ISO-2), nonché il numero di passeggeri e il numero di voli registrati per quella specifica tratta.

```

1 { "states": [
2   {
3     "name": "Bahrain",
4     "iso_code": "BH",
5     "continent": "Asia"
6   },
7   {
8     "name": "Uruguay",
9     "iso_code": "UY",
10    "continent": "South America"
11  },
12  {
13    "name": "Arizona",
14    "iso_code": "US-AZ",
15    "continent": "North America"
16  },
17  ...
18 ]
19
20 "links": [
21   {
22     "US_state": "New York",
23     "FG_state": "Bahrain",
24     "flights": 36,
25     "passengers": 4738,
26     "year": 1994,
27     "month": 7
28   },
29   {
30     "US_state": "Arizona",
31     "FG_state": "Canada",
32     "flights": 262,
33     "passengers": 25773,
34     "year": 2003,
35     "month": 4
36   },
37   ...
38 ]
39 }

```

2.2.3 Definizioni dei tipi di Dataset e degli Attributi

Tipi di Dataset:

- **Rete** → Gli elementi (nodi) rappresentano gli Stati coinvolti nel traffico aereo, mentre gli archi rappresentano i collegamenti aerei tra gli Stati.
[Cardinalità: nodi = 151 Stati; collegamenti = 139650]
- **Geometria** → Gli elementi rappresentano gli Stati su una mappa geografica.
[Cardinalità: 226 Stati]

Tipi di Attributi per i nodi:

- **Nome dello Stato:** attributo categorico [Cardinalità: 151 Stati].
- **Codice ISO:** attributo categorico [Cardinalità: 151 codici ISO].
- **Continente:** attributo categorico [Cardinalità: 7 continenti].

Tipi di Attributi per i collegamenti:

- **Anno:** attributo ordinale [Cardinalità: intervallo da 1990 a 2020 = 31].
- **Mese:** attributo ordinale [Cardinalità: 12].
- **Stato US:** attributo categorico [Cardinalità: 49 Stati U.S.].
- **Stato estero:** attributo categorico [Cardinalità: 102 stati esteri].
- **Numero di Passeggeri:** attributo quantitativo [Cardinalità: Intervallo da 1 a 840185].
- **Numero di Voli:** attributo quantitativo [Cardinalità: Intervallo da 1 a 8099].

2.3 Task Modeling

1. Trovare uno Stato.
2. Trovare gli Stati US più connessi (con più collegamenti) in uno specifico mese e anno.
3. Determinare se due Stati hanno un collegamento diretto, in uno specifico mese e anno.
4. Trovare collegamenti di uno specifico Stato US verso l'estero in uno specifico mese e anno.
5. Trovare collegamenti più trafficati dai passeggeri in uno specifico mese e anno.
6. Trovare collegamenti più trafficati da voli in uno specifico mese e anno.
7. Determinare se due Stati hanno una destinazione raggiungibile in comune in uno specifico mese e anno.
8. Trovare gli Stati che si trovano nello stesso continente.
9. Determinare se ci sono molti collegamenti tra gli Stati US e gli Stati del continente X in uno specifico mese e anno.

Definiamo ogni task come coppia {azione, target}:

1. Trovare uno Stato.
→ lookup del nodo (Stato).
2. Trovare gli Stati US più connessi (con più collegamenti) in uno specifico mese e anno.
→ filtering delle tuple che si riferiscono al mese e anno selezionati.
→ locate degli Stati con valore rappresentante il numero di collegamenti più elevato.
→ compare dei valori degli Stati.
3. Determinare se due Stati hanno un collegamento diretto, in uno specifico mese e anno.
→ filtering delle tuple che si riferiscono al mese e anno selezionati.
→ identify del collegamento tra i due nodi (Stati).
→ lookup dei due nodi (Stati).

4. Trovare collegamenti di uno specifico Stato US verso l'estero in uno specifico mese e anno.
 - **filtering** delle tuple che si riferiscono al mese e anno selezionati.
 - **identify** dei collegamenti tra il nodo dello Stato US e gli Stati esteri.
 - **lookup** dello Stato.
 - **explore** degli Stati connessi.
5. Trovare collegamenti più trafficati dai passeggeri in uno specifico mese e anno.
 - **filtering** delle tuple che si riferiscono al mese e anno selezionati.
 - **locate** dei collegamenti con maggior numero di passeggeri che lo trafficano.
 - **compare** del numero di passeggeri che trafficano i collegamenti.
6. Trovare collegamenti più trafficati da voli in uno specifico mese e anno.
 - **filtering** delle tuple che si riferiscono al mese e anno selezionati.
 - **locate** dei collegamenti con maggior numero di voli che lo trafficano.
 - **compare** del numero di voli che trafficano i collegamenti.
7. Determinare se due Stati hanno una destinazione raggiungibile in comune in uno specifico mese e anno.
 - **filtering** delle tuple che si riferiscono al mese e anno selezionati.
 - **identify** di percorsi di lunghezza 2 tra i due nodi (Stati).
 - **lookup** dei due nodi (Stati).
8. Trovare gli Stati che si trovano nello stesso continente.
 - **browse** degli Stati.
 - **compare** dell'attributo continente.
9. Determinare se ci sono molti collegamenti tra gli Stati US e gli Stati del continente X in uno specifico mese e anno.
 - **filtering** delle tuple che si riferiscono al mese e anno selezionati.
 - **identify** dei collegamenti tra i nodi (Stati) US e i nodi del continente X.
 - **browse** dei nodi (Stati) US e dei nodi del continente X.

3. Progettazione

3.1 Progettazione della Visualizzazione

3.1.1 Choropleth Map

La Choropleth Map permette di visualizzare immediatamente il numero di collegamenti per ciascuno Stato US.

Nei task richiesti, comprendere le relazioni spaziali tra i vari Stati è di primaria importanza, la scelta di utilizzare una Choropleth Map permette di analizzare tali relazioni in modo facile ed immediato.

D'altra parte, la Choropleth Map, pur facilitando l'analisi visiva, presenta un limite significativo: l'evidenza visiva è influenzata dalla dimensione della regione, non dalla reale importanza rispetto al valore dell'attributo.

Tuttavia, questo limite può essere trascurato, poiché il vantaggio di una comprensione facile e immediata risulta più importante per l'adempimento del task primario.

Marks

- Items: interlocking marks (Stati)

Channels

- Posizione, area e forma per rappresentare una regione geografica.
- Colore per rappresentare l'attributo quantitativo (numero di collegamenti totali).

Semantics

Nella choropleth map gli interlocking marks rappresentano gli Stati, mentre, il colore (utilizzando una sequential colormap) rappresenta il numero dei collegamenti per un determinato Stato US.

3.1.2 Node-Link Connection Map

La connection Map permette di visualizzare in modo intuitivo i collegamenti tra gli Stati US e gli Stati esteri, fornendo una chiara visione delle relazioni geografiche tra le diverse aree.

I nodi, rappresentati dagli interlocking marks degli Stati, sono posizionati su punti fissi della mappa, corrispondenti ai centroidi degli Stati stessi.

Mentre i link, rappresentano il collegamento tra due Stati. In aggiunta il colore del link incorpora l'informazione riguardante il numero di passeggeri o il numero di voli (a scelta dell'utente).

Tuttavia un limite di questa rappresentazione è che, nelle aree con un elevato numero di collegamenti, le linee possono sovrapporsi, riducendo la leggibilità.

Per questo motivo si è scelto di limitare la visualizzazione ad un determinato mese e anno specifico; permettendo, inoltre, all'utente di selezionare gli Stati US di cui vuole vedere i collegamenti.

Marks

- Items (nodi): interlocking marks (Stati).
- Links: collegamenti (segmenti).

Channels

- **Channels per gli items (nodi):** attributo categorico (continente): color hue dell'interlocking mark.
- **Channels per i links:** attributo quantitativo (numero di passeggeri/voli): colore (sequential colormap).

Semantics

Gli interlocking marks (Stati) rappresentano i nodi della rete, mentre, i links rappresentano i collegamenti tra gli Stati.

3.2 Progettazione dell'interazione

Operazioni di Interazione

- **Selection:** selezione di uno o più Stati US di interesse.
- **Exploration:** panning.
- **Zooming**
- **Details-on-demand:** visualizzazione di etichette informative:
 - Stati US: nome e numero di collegamenti
 - Stati esteri: nome
 - Collegamenti: Stato US, Stato estero, numero di passeggeri e numero di voli.
- **Reconfiguring:** selezione attributo visualizzato per i collegamenti attraverso il colore (numero di passeggeri/voli).
- **Encoding:** transizione tra gli idiomi visuali (Choropleth map e Connection Map).
- **Filtering:** query dinamiche in base a mese, anno e Stato selezionato.

Paradigma di interazione

Interazione incrementale: l'esplorazione inizia da una piccola porzione della rete, limitata agli Stati US. Successivamente, la visualizzazione si espande su richiesta, mostrando i collegamenti specificati dall'utente.

3.3 Scelte Tecnologiche ed Architetture

Tipologia di sistema e tecnologie utilizzate

Il sistema progettato è una Web Application, la quale fornisce una rappresentazione interattiva della rete di collegamenti, permettendo agli utenti di esplorare i dati attraverso visualizzazioni dinamiche.

Per lo sviluppo dell'applicazione sono state utilizzate le seguenti tecnologie:

- **HTML:** utilizzato per definire la struttura e organizzare il contenuto della pagina web.
- **CSS:** impiegato per definire lo stile grafico e garantire un design responsive.

- **JavaScript:** adottato per implementare la logica applicativa e gestire l'interattività della visualizzazione.
- **Libreria D3.js:** utilizzata per creare e gestire la visualizzazione interattiva.

Strumenti e tecniche di interazione

Per garantire un'esperienza utente completa, vengono adottate le seguenti tecniche e strumenti di interazione, basate sull'utilizzo del mouse:

- **Selection** → ogni singolo Stato US di interesse può essere selezionato utilizzando il pulsante sinistro del mouse.
- **Exploration** → l'utente può spostarsi all'interno della mappa geografica tenendo premuto il tasto sinistro del mouse e trascinandolo, per poi rilasciarlo quando si arriva nella zona di interesse.
- **Zooming** → il livello di zoom può essere adattato utilizzando la rotella di scorrimento del mouse.
- **Details-on-demand** → l'utente può visualizzare le etichette informative passando il cursore sugli Stati o sui collegamenti.
- **Reconfiguring** → attraverso dei pulsanti, l'utente può selezionare se il colore dei collegamenti rappresenta il numero di passeggeri o il numero di voli.
- **Encoding:** → una volta selezionato uno o più Stati US di interesse, la Choropleth Map non sarà più visibile e saranno visualizzati solamente i collegamenti relativi agli Stati US selezionati, passando così alla visualizzazione della Connection Map.
- **Filtering:** → l'utente può selezionare il mese e l'anno di interesse utilizzando due slider dedicati. Inoltre, in base alla selezione degli Stati US, verranno visualizzati esclusivamente i collegamenti ad essi associati.

4. Realizzazione

All'avvio del sistema viene visualizzata la mappa, eseguendo uno zoom-in sugli Stati Uniti.

Sugli Stati US viene visualizzata una choropleth map rappresentante il numero di collegamenti verso l'estero di ogni Stato (ovvero il suo grado) nello specifico mese e anno selezionato (vedi figura 4.1). L'utilizzo di una scala logaritmica per la colormap sequenziale è stato adottato per rendere più distinguibili i valori più piccoli, che altrimenti risultano molto simili con una gamma lineare dominata dai valori elevati.

Sul lato destro dello schermo si trovano gli slider che permettono di selezionare il mese e anno di interesse.

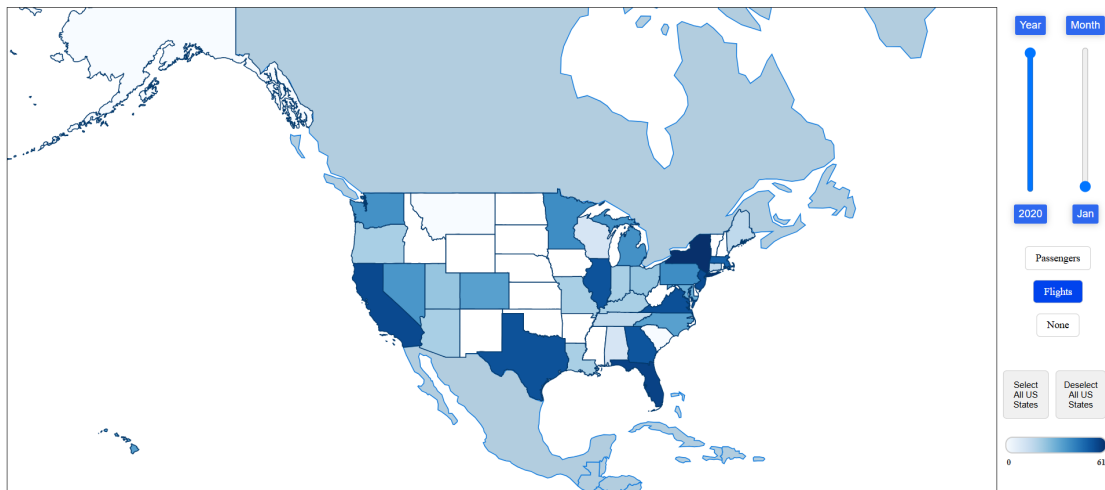


Figura 4.1: Schermata iniziale con zoom-in sugli Stati Uniti e choropleth map mostrata.

L'interfaccia permette all'utente di selezionare gli Stati US di cui vuole vedere i collegamenti disponibili attraverso il tasto sinistro del mouse; a questo punto sulla

mappa scompare la choropleth map e, a seguito di uno zoom-out, compaiono gli archi rappresentanti i collegamenti degli Stati selezionati (vedi figura 4.2).



Figura 4.2: Schermata dopo la selezione del Texas.

È possibile la scelta multipla degli Stati. Sul lato destro dello schermo sono presenti due pulsanti che permettono di selezionare tutti gli Stati US o deselegionare tutti gli Stati attualmente evidenziati in modo efficiente.

In modo predefinito gli archi rappresentanti i collegamenti vengono colorati utilizzando una colormap sequenziale, anche in questo caso logaritmica (per gli stessi motivi evidenziati precedentemente), che descrive il numero di passeggeri che hanno trafficato il rispettivo collegamento nel mese e anno selezionato. Sul lato destro dell'interfaccia sono disponibili dei pulsanti che permettono di cambiare l'informazione codificata dal colore degli archi: oltre al numero di passeggeri, si può scegliere di visualizzare il numero dei voli, oppure semplicemente di non utilizzare questo channel, colorando tutti gli archi di rosso.

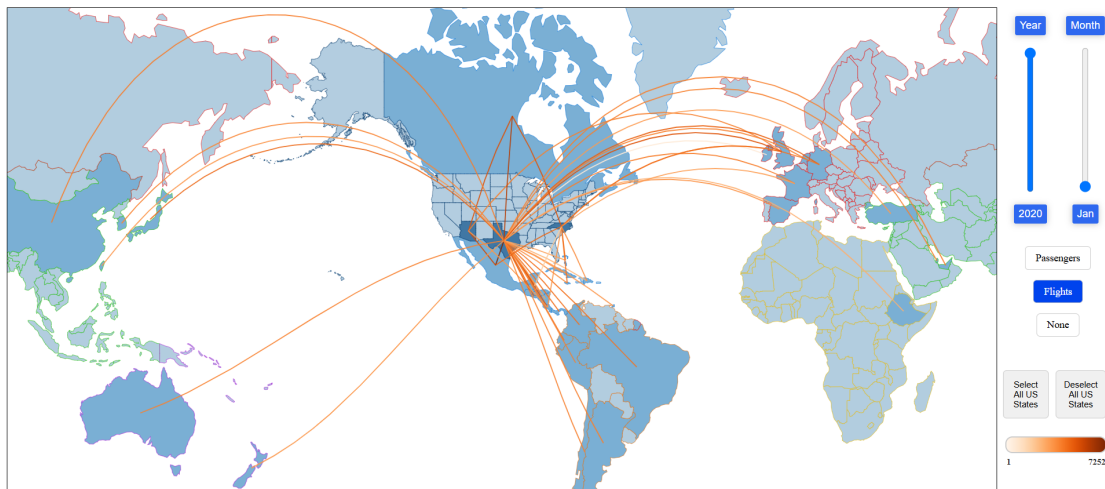


Figura 4.3: Esempio di selezione multipla: in questo caso sono stati selezionati gli Stati del Texas, Arizona e Nord Carolina. Inoltre è stato selezionato il numero di voli come attributo codificato dal colore degli archi.

La mappa è esplorabile dall'utente attraverso zoom e panning. Inoltre, l'utente può visualizzare informazioni aggiuntive attraverso delle tooltips visibili a seguito di hovering sugli Stati ed i collegamenti.

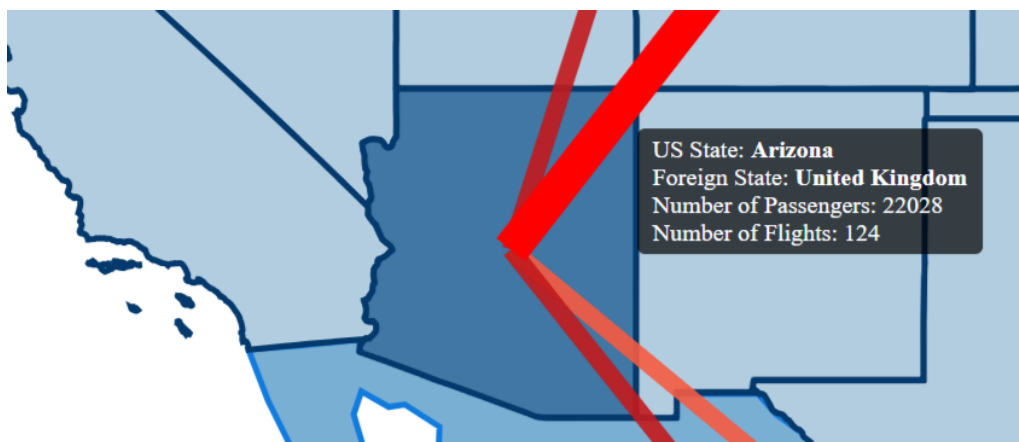


Figura 4.4: Esempio di hovering effettuato sul collegamento tra Arizona e UK.

La realizzazione completa del progetto, inclusi il codice sorgente e i dataset utilizzati, è disponibile pubblicamente nella repository GitHub al seguente link: <https://github.com/RiccardoN001/us-airTraffic-analytics.git>.