

# LOTNISKO CHOPINA WARSZAWA

## ANALIZA SIECI POŁĄCZEŃ W ROKU 2018

Adam Nowak

Uniwersytet Warszawski

Warszawa

04.07.2024

## Spis treści

Spis oznaczeń .....	3
1. Cel projektu .....	4
2. Przegląd danych .....	4
3. Dane .....	5
4. Przygotowanie danych do programu .....	5
5. Główny program .....	6
5.1 Biblioteki .....	6
5.2 Wczytanie danych .....	6
5.3 Mapowanie danych .....	6
5.4 Sprawdzenie danych.....	7
5.5 Podział roku.....	7
6. Analiza odlotów i przylotów z Lotniska Chopina .....	8
6.1 Wizualizacja wyników .....	9
7. Wnioski .....	17
8. Perspektywy dalszego rozwoju projektu.....	18
9. Spis ilustracji i tabel .....	19

## Spis oznaczeń

*Id* – unikalny identyfikator powiązany z każdym lotem (FlightAware flight id).

*Ident* – znak wywoławczy identyfikujący lot. Kod linii lotniczej ICAO.

*Reg* – numer rejestracyjny statku powietrznego.

*Orig* – kod lotniska ICAO, miejsce rozpoczęcia lotu.

*Dest* – kod lotniska ICAO, miejsce zakończenia lotu.

*Pitr* - wartość znacznika czasu.

# 1. Cel projektu

Celem projektu była analiza siatki połączeń lotniska Chopina w Warszawie. Do zadania wykorzystano bazę danych „fsig-raw” z 2018 roku dostępną na serwerze fsig-students.ocean.icm.edu.pl.

Program został napisany w języku Python ze względu na dużą ilość danych. Na serwerze użyto wersji 2.7, natomiast po uzyskaniu danych lokalnie wykorzystano wersję 3.11 oraz środowisko Jupyter Notebook.

## 2. Przegląd danych

Początkowo skupiono się na przeanalizowaniu dostępnych danych. Każdy zestaw zawiera kolejno pola: orig, dest, pitr, type, ident, clock, lon, lat oraz inne istotne zmienne (Rysunek 1).

```
> db.rawFaSignals.find().limit(2);
{ "_id" : ObjectId("5a497a04870c008695942e4d"), "key" : "1514764800#SJY590#72592
fc3bcb6e36761fb4982e3bda1dc#v2", "value" : { "pitr" : "1514764800", "type" : "position", "ident" : "
SJY590", "alt" : "5300", "clock" : "1514764792", "facility_hash" : "910
f66749f1ce9655b9f41ce7850e239974cd798", "facility_name" : "FlightAware ADS-B", "id" : "SJY590
-1514584800-schedule-0001", "gs" : "213", "heading" : "30", "hexid" : "8A0670", "lat" : "-5.33489", "
lon" : "119.40396", "updateType" : "A", "altChange" : "D", "air_ground" : "A" } }
```

*Rysunek 1 Przegląd danych (dla typu position)*

Z zestawu o parametrze type: „Flightplan” (Rysunek 2, Rysunek 3) wyszukano lotnisko Chopina po kodzie ICAO – „EPWA” oraz zapisano do pliku interesujące dane .

```
# Kryteria wyszukiwania dokumentów w określonym przedziale czasowym
date_criteria = {
    "$and": [
        {"value.pitr": {"$gte": str(start_epoch)}},
        {"value.pitr": {"$lt": str(end_epoch)}},
        {"value.orig": "EPWA"}
    ]
}
```

*Rysunek 2 Wycinek kodu, kryteria wyszukiwania połączeń z lotniska Chopina*

```
# Kryteria wyszukiwania dokumentów w określonym przedziale czasowym
date_criteria = {
    "$and": [
        {"value.pitr": {"$gte": str(start_epoch)}},
        {"value.pitr": {"$lt": str(end_epoch)}},
        {"value.dest": "EPWA"}
    ]
}
```

*Rysunek 3 Wycinek kodu, kryteria wyszukiwania połączeń do lotniska Chopina*

## 3. Dane

Dane zebrano dla każdego miesiąca w roku 2018. Początkowo planowano zebrać dane dla wszystkich lotnisk w Polsce oraz dla kolejnych lat, jednak ze względu na długi czas odczytu oraz problemy z kopiowaniem danych na komputer lokalny, ostatecznie zebrano dane tylko za jeden rok dla lotniska Chopina. Czas obliczeń dla jednego dnia został pokazany na Rysunek 4. W ten sposób otrzymano 24 pliki JSON z przylotami oraz odlotami z lotniska Chopina.

```
Czas połączenia z bazą danych: 0.00 sekund
Czas zapytania do bazy danych: 127.26 sekund
Czas przetwarzania danych: 0.02 sekund
Czas zapisu danych do pliku: 0.01 sekund
Czas odczytu danych z pliku: 0.01 sekund
Całkowity czas wykonania: 135.05 sekund
```

Rysunek 4 Czas wyszukiwania danych dla 1 stycznia 2018 roku

```
{"ident": "ELY562", "dest": "LLBG", "pitr": "2018-01-01 01:00:47", "id": "ELY562-1514528788-airline-0039", "reg": "4XEKA", "orig": "EPWA"},
{"ident": "LOT79", "dest": "RJAA", "pitr": "2018-01-01 01:11:59", "id": "LOT79-1514528788-airline-0032", "reg": "SPLRG", "orig": "EPWA"},
{"ident": "QTR262", "dest": "OTHH", "pitr": "2018-01-01 01:12:46", "id": "QTR262-1514528789-airline-0314", "reg": "A7AHQ", "orig": "EPWA"},
{"ident": "BAW847", "dest": "EGLL", "pitr": "2018-01-01 02:19:49", "id": "BAW847-1514615190-airline-0246", "reg": "GEUUM", "orig": "EPWA"},
{"ident": "BAW851", "dest": "EGLL", "pitr": "2018-01-01 05:24:49", "id": "BAW851-1514615190-airline-0120", "reg": "GEUYX", "orig": "EPWA"},
{"ident": "AFR1247", "dest": "LPPG", "pitr": "2018-01-01 06:34:14", "id": "AFR1247-1514612100-schedule-0000", "reg": "FGTAU", "orig": "EPWA"}
```

Rysunek 5 Dane zawierające połączenia do lotniska Chopina

```
{"ident": "CCA737", "dest": "EPWA", "pitr": "2018-01-01 01:07:24", "id": "CCA737-1514577000-schedule-0000", "reg": "B6131", "orig": "ZBAA"},
{"ident": "BAW850", "dest": "EPWA", "pitr": "2018-01-01 02:05:36", "id": "BAW850-1514615190-airline-0172", "reg": "GEUYX", "orig": "EGLL"},
{"ident": "LOT92", "dest": "EPWA", "pitr": "2018-01-01 02:18:06", "id": "LOT92-1514594700-schedule-0000", "reg": "SPLRG", "orig": "ZBAA"},
{"ident": "LOT80", "dest": "EPWA", "pitr": "2018-01-01 03:00:31", "id": "LOT80-1514598900-schedule-0002", "reg": "SPLRG", "orig": "RJAA"},
{"ident": "ELY563", "dest": "EPWA", "pitr": "2018-01-01 05:10:56", "id": "ELY563-1514607000-schedule-0000", "reg": "4XEHE", "orig": "LLBG"},
{"ident": "UAE179", "dest": "EPWA", "pitr": "2018-01-01 05:12:36", "id": "UAE179-1514606700-schedule-0000", "reg": "A6EPA", "orig": "OMDB"}
```

Rysunek 6 Dane zawierające połączenia z lotniska Chopina

Ścieżka do programów na github:

1. `ONS/program_bash/origin_chopin.py`
2. `ONS/program_bash/destination_chopin.py`

## 4. Przygotowanie danych do programu

Przed rozpoczęciem analizy konieczne było zmodyfikowanie danych, ponieważ zawierały one przypadki, które nie nadawały się do analizy. Napotkano problemy z polami „dest” oraz „ident”, gdzie wartość nie była czteroznakowym kodem lub zawierała nieznany numer ICAO. W tym celu opracowano program, który modyfikował dane.

```
"dest": "L 47.41730 8.53003",
unwanted_orig = ['LFAV', 'LFTB', 'LSGL'...
```

Rysunek 7 Problematiczne dane

W celu zidentyfikowania kodów oraz połączenia ich z państwami, regionami i kontynentami, wykorzystano dostępne w Internecie pliki zawierające niezbędne dane.

1. `iata-icao.csv` (<https://github.com/ip2location/ip2location-iata-icao>),
2. `all.csv` (<https://github.com/luke/ISO-3166-Countries-with-Regional-Codes/blob/master/all/all.csv>).

Do pierwszego pliku dodano nowe zmienne, aby uzupełnić brakujące kody ICAO. W drugim pliku zmodyfikowano nazwy niektórych państw. Plikom nadano nowe nazwy.

Program do obróbki danych znajduje się w folderze: `ONS/program/data_procces.ipynb`.

## 5. Główny program

Program składa się z części wczytania bibliotek, załadowania danych, mapowania danych, analizy danych oraz wizualizacji wyników.

Program do analizy danych dla siatki połączeń lotniska Chopina:

1. *ONS/program/chopin\_analysis\_origin.ipynb*
2. *ONS/program/chopin\_analysis\_destination.ipynb*

### 5.1 Biblioteki

- JSON - praca z danymi w formacie JSON (wczytywanie i zapisywanie)
- OS - operacje na plikach i ścieżkach (sprawdzanie istnienia plików)
- RE – manipulacja tekstem i wyrażeń regularnych
- NUMPY – wykorzystana do obliczeń
- PANDAS – przetwarzanie oraz analiza danych
- MATPLOTLIB.PYPLLOT – tworzenie wykresów oraz wizualizacja danych

### 5.2 Wczytanie danych

Dane zostały wczytane z poprzednio utworzonych plików zawierających loty dla kolejnych miesięcy.

```
import json
import os
import re
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

Rysunek 8 Użyte biblioteki

```
file_list_orig = ['1.json', '2.json', '3.json' ...]
```

Rysunek 9 Wczytanie danych

### 5.3 Mapowanie danych

Aby przeprowadzić analizę danych należało połączyć wpierw kody ICAO z kodami państw. Wykorzystano do tego plik „combined\_iata\_icao.csv”. Następnie wykorzystując plik „all1” zmapowano dwuznakowe kody państw na państwa, kontynenty oraz regiony na świecie. Następnie dane zostały zapisane w postaci tabelarycznej.

```
# Mapowanie
icao_to_country_code = airports_df.set_index('icao')['country_code'].to_dict()
country_code_to_name = countries_df.set_index('alpha-2')['name'].to_dict()
country_code_to_continent = countries_df.set_index('alpha-2')['region'].to_dict()
country_code_to_region = countries_df.set_index('alpha-2')['sub-region'].to_dict()
```

Rysunek 10 Mapowanie danych

## 5.4 Sprawdzenie danych

Po wczytaniu danych sprawdzono czy wszystkie kody ICAO posiadają swoje odniesienia do państw. Napotkano parę problemów zarówno w przylotach jak i odlotach, przez co wykorzystano dodatkowa obróbkę danych o której wspomniano w rozdziale 4. Przygotowanie danych do programu.

```
# Znalezienie niezmapowanych kodów ICAO
unmapped_icao = []

for icao_code in df['dest_country_code'].unique():
    if icao_code not in country_code_to_name:
        unmapped_icao.append(icao_code)

# Wypisanie niezmapowanych kodów ICAO
print("Niezmapowane kody ICAO:")
for icao_code in unmapped_icao:
    print(icao_code)
```

Rysunek 11 Sprawdzenie poprawności mapowania

## 5.5 Podział roku

Ostatnim krokiem przed rozpoczęciem analizy było dodanie pór roku.

```
season_map = {
    1: 'Zima', 2: 'Zima', 3: 'Wiosna',
    4: 'Wiosna', 5: 'Wiosna', 6: 'Wiosna',
    7: 'Lato', 8: 'Lato', 9: 'Lato',
    10: 'Jesień', 11: 'Jesień', 12: 'Zima'
}
df['season'] = df['month'].map(season_map)
```

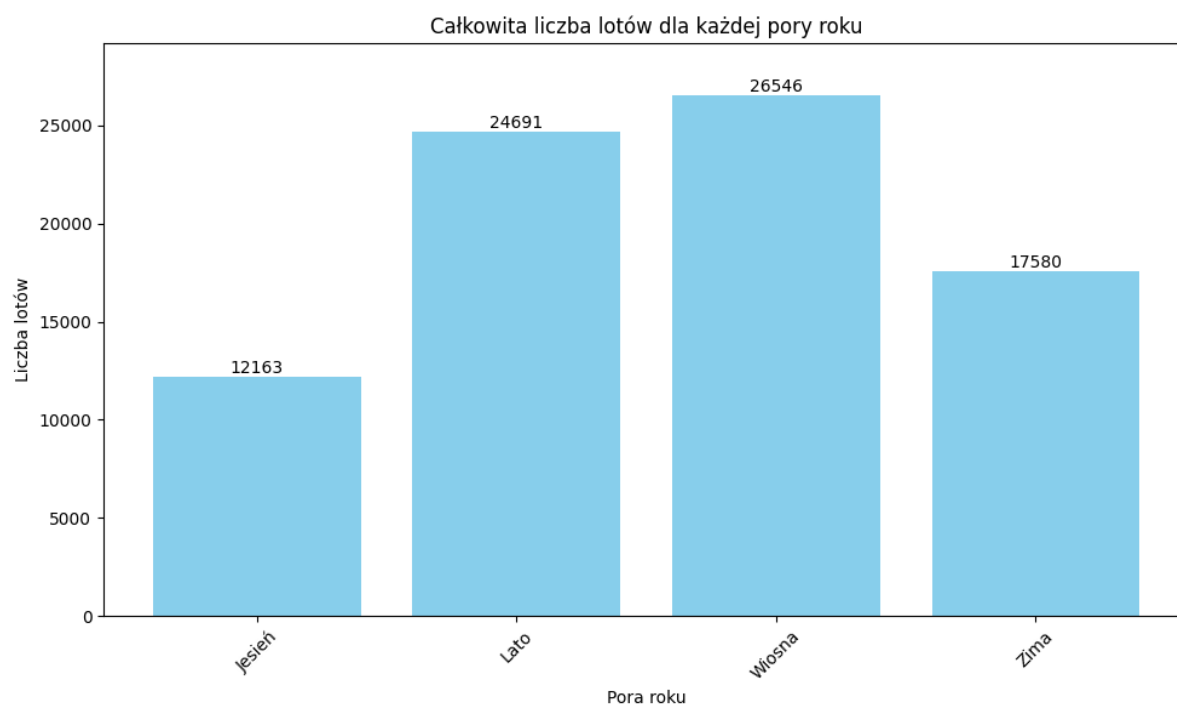
Rysunek 12 Podział miesięcy na pory roku

## 6. Analiza odlotów i przylotów z Lotniska Chopina

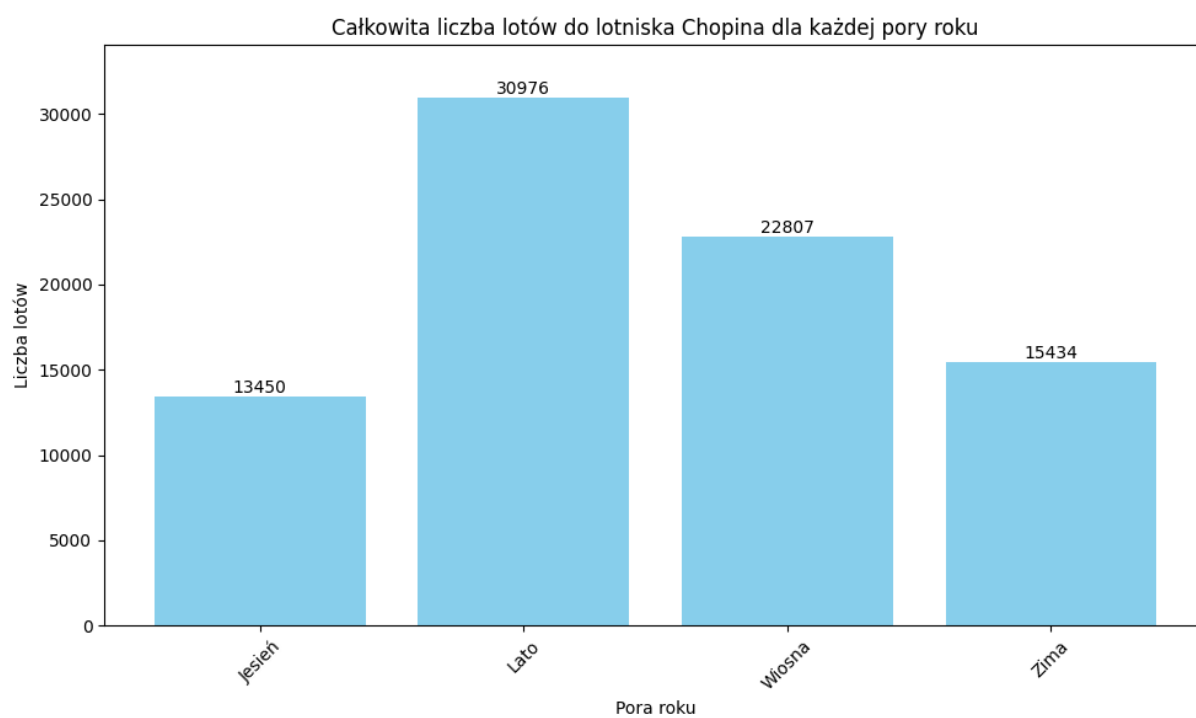
- **Analiza lotów w zależności od pory roku:**
  - Analiza liczby lotów odlotowych w różnych porach roku.
  - Analiza liczby lotów przylotowych w różnych porach roku.
- **Analiza lotów do/z kontynentów w zależności od pory roku:**
  - Analiza lotów odlotowych do różnych kontynentów w zależności od pory roku.
  - Analiza lotów przylotowych z różnych kontynentów w zależności od pory roku.
- **Analiza najpopularniejszych kierunków/regionów w zależności od pory roku:**
  - Analiza najpopularniejszych kierunków odlotowych w różnych porach roku.
  - Analiza najpopularniejszych regionów, z których podróżuje się do lotniska Chopina, w różnych porach roku.
- **Analiza lotów dla 10 najpopularniejszych państw w ciągu roku:**
  - Analiza lotów odlotowych dla 10 najpopularniejszych państw w ciągu roku.
  - Analiza lotów przylotowych dla 10 najpopularniejszych połączeń z państw w ciągu roku.
- **Analiza lotów dla 10 najpopularniejszych państw w Europie w zależności od pory roku:**
  - Analiza lotów odlotowych dla 10 najpopularniejszych państw w Europie w zależności od pory roku.
  - Analiza lotów przylotowych dla 10 najpopularniejszych połączeń z państw w Europie w zależności od pory roku.
- **Analiza lotów dla 10 najpopularniejszych państw poza Europą w zależności od pory roku:**
  - Analiza lotów odlotowych dla 10 najpopularniejszych państw poza Europą w zależności od pory roku.
  - Analiza lotów przylotowych dla 10 najpopularniejszych połączeń z państw spoza Europy w zależności od pory roku.
- **Analiza przepustowości lotniska w zależności od miesiąca:**
  - Analiza przepustowości lotniska pod kątem odlotów w zależności od miesiąca.
  - Analiza przepustowości lotniska pod kątem przylotów w zależności od miesiąca.
- **Analiza przepustowości lotniska w zależności od pory roku:**
  - Analiza przepustowości lotniska pod kątem odlotów w zależności od pory roku.
  - Analiza przepustowości lotniska pod kątem przylotów w zależności od pory roku.



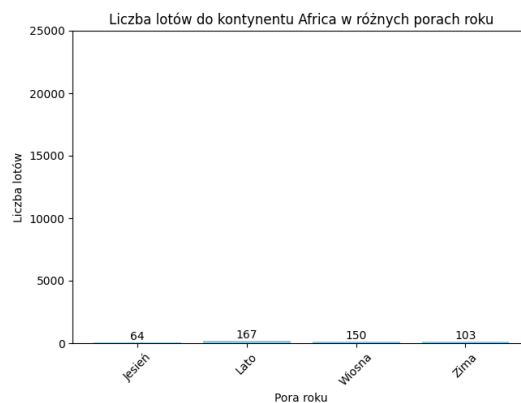
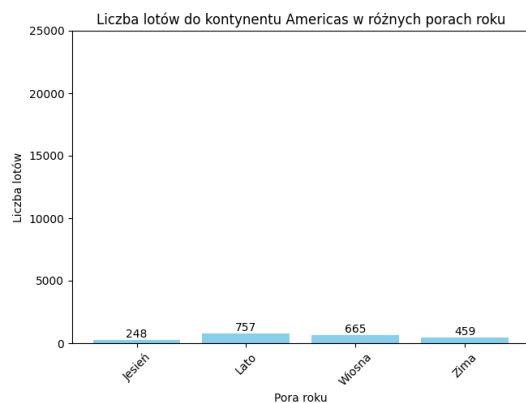
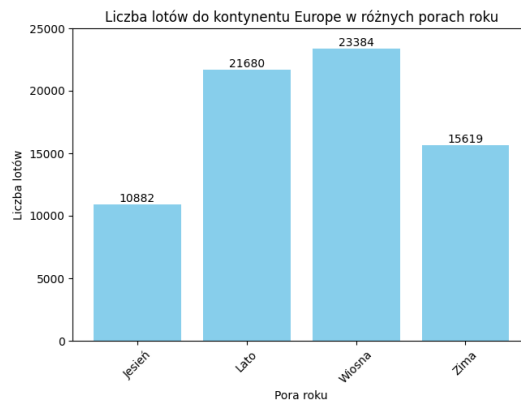
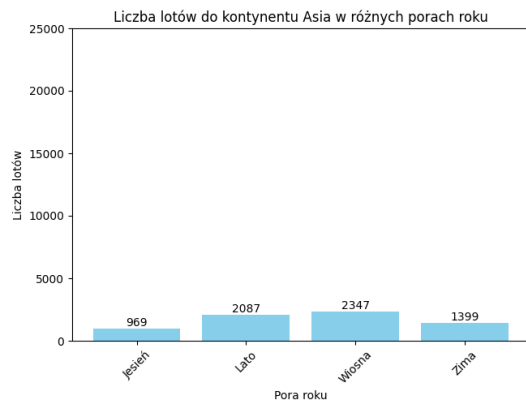
## 6.1 Wizualizacja wyników



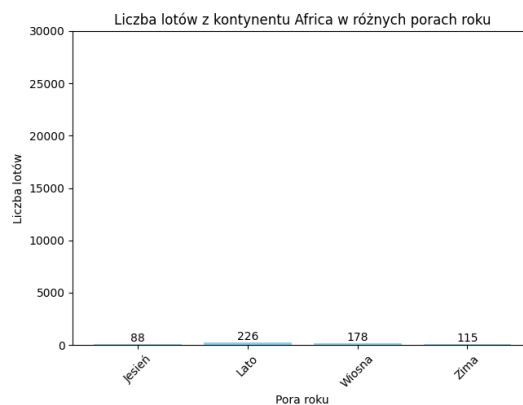
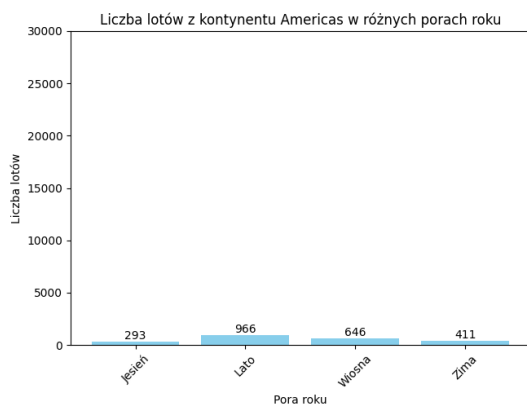
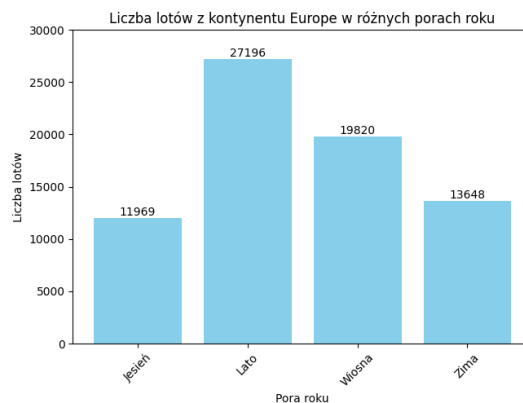
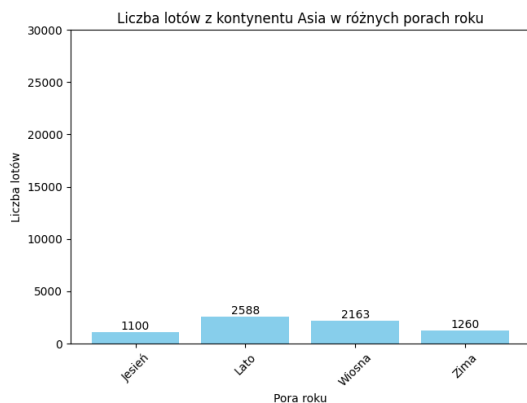
Rysunek 13 Liczba odlotów z lotniska Chopina w ciągu roku



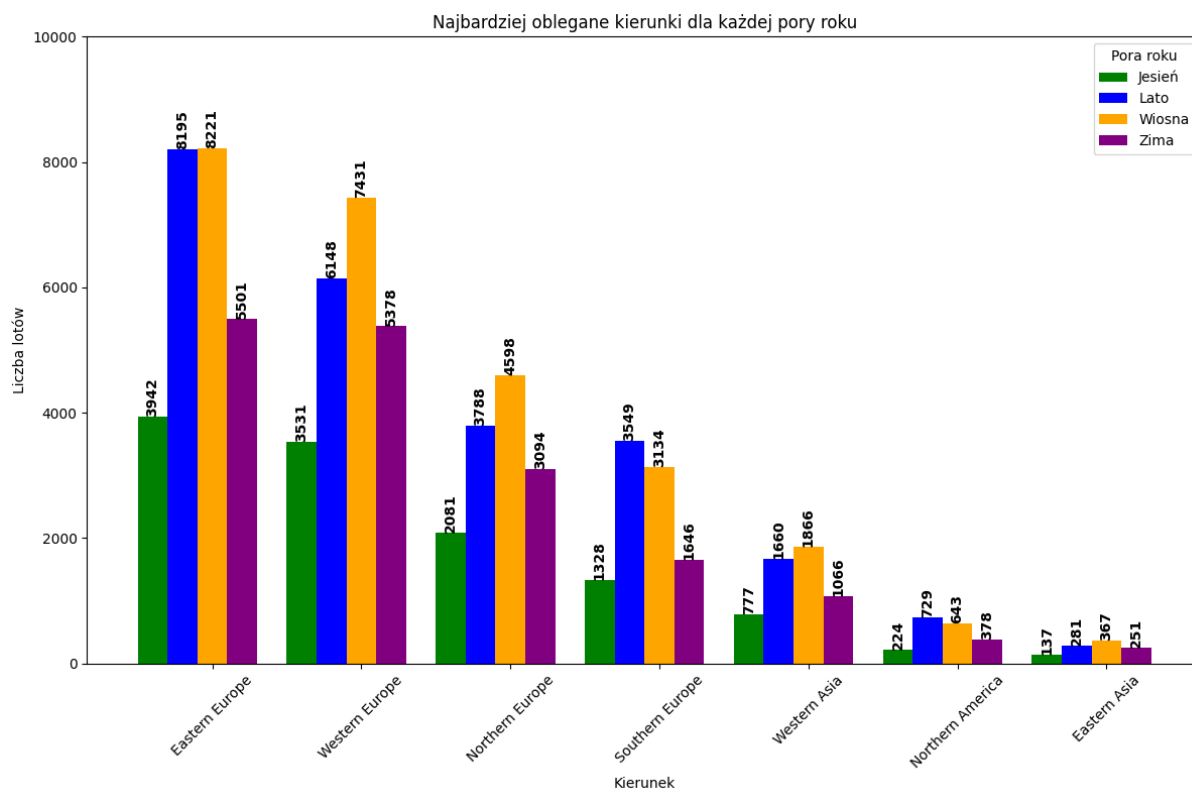
Rysunek 14 Liczba przylotów do lotniska Chopina w ciągu roku



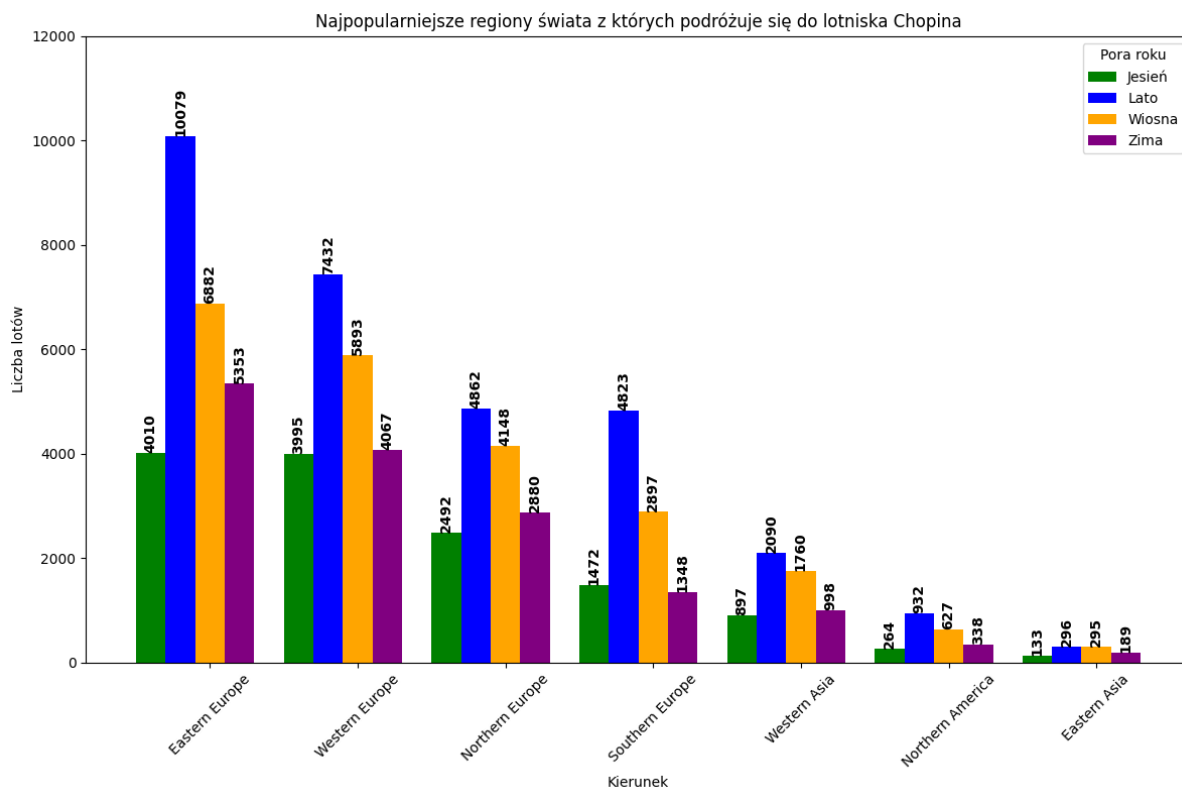
Rysunek 15 Liczba lotów do kontynentów w zależności od pory roku



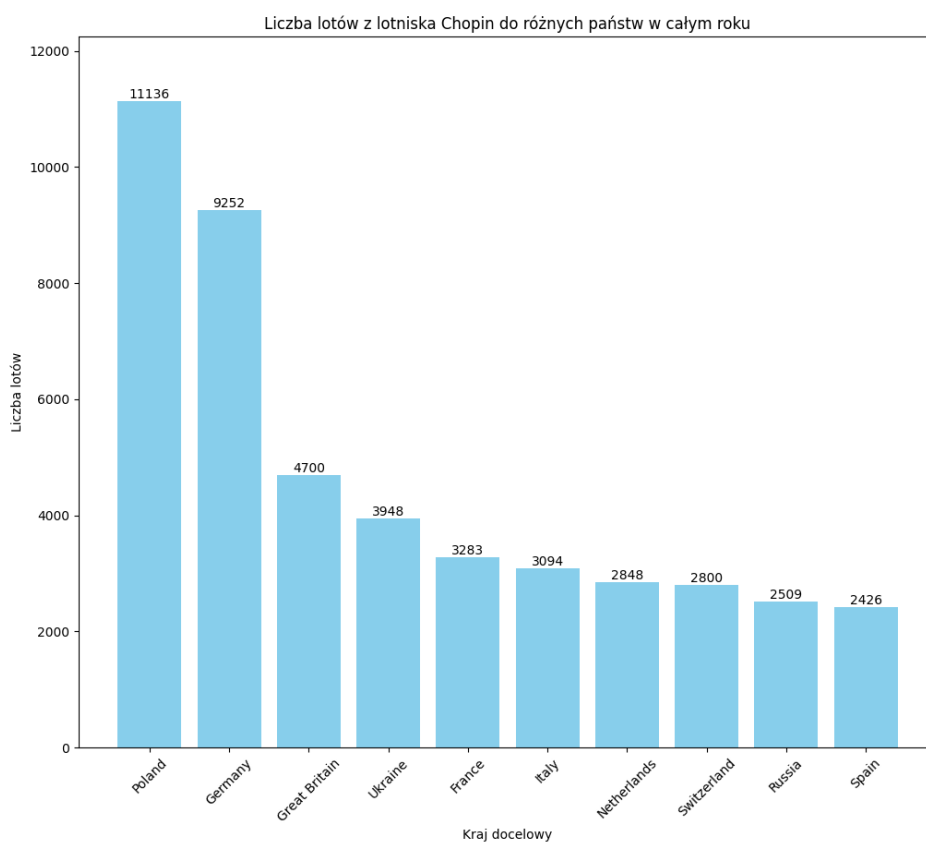
Rysunek 16 Liczba przylotów z kontynentów do lotniska Chopina w zależności od pory roku



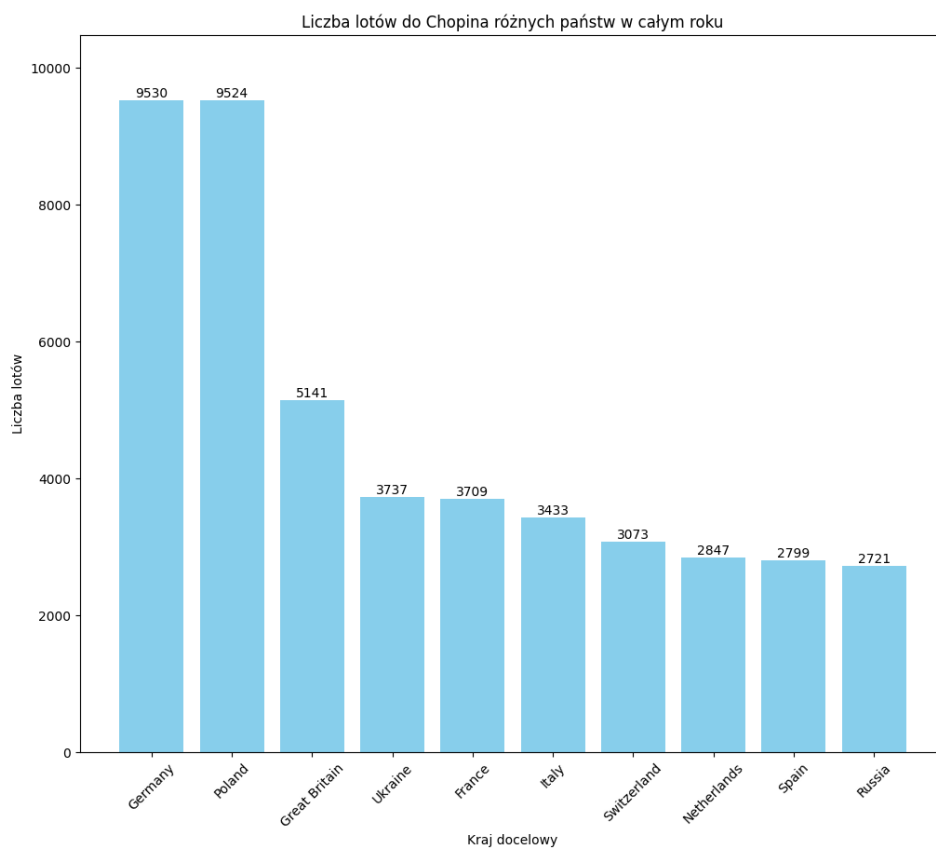
Rysunek 17 Najpopularniejsze kierunki w zależności od pory roku



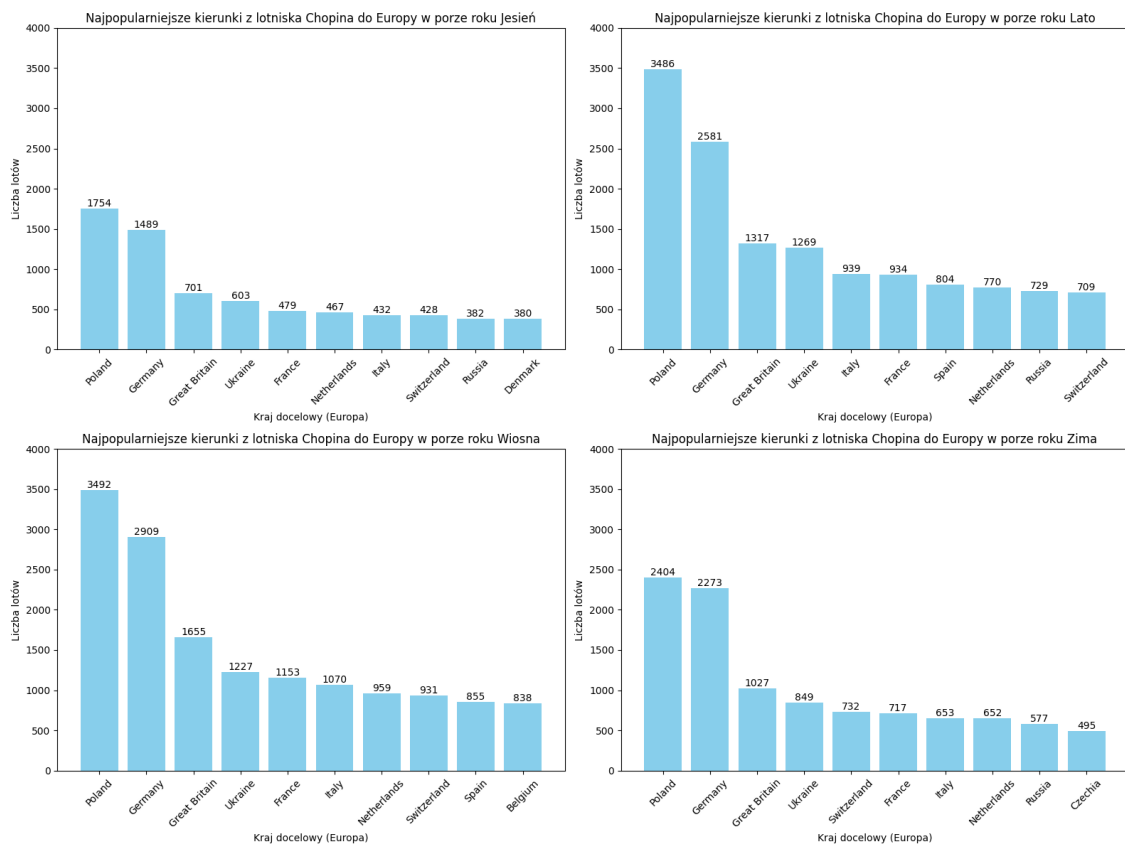
Rysunek 18 Najpopularniejsze regiony świata, z których podróżuje się do lotniska Chopina w zależności od pory roku



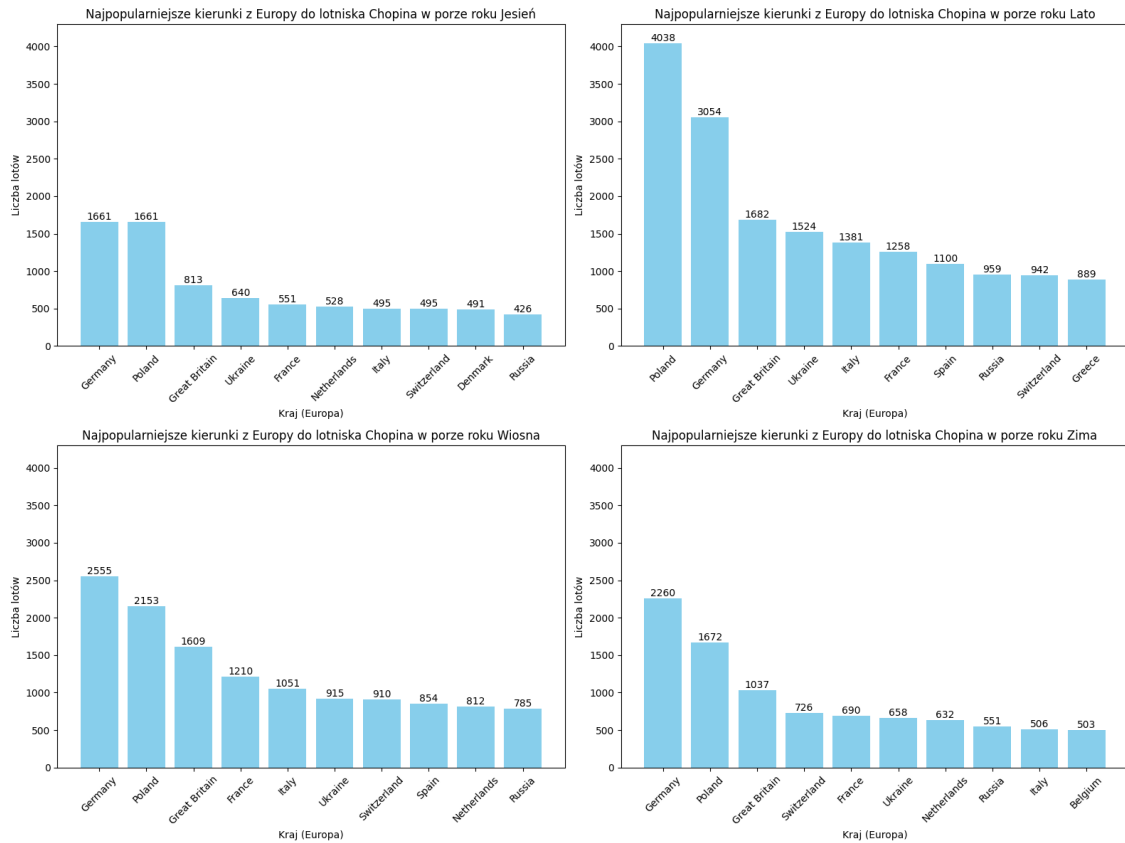
Rysunek 19 Liczba lotów do dziesięciu najpopularniejszych państw



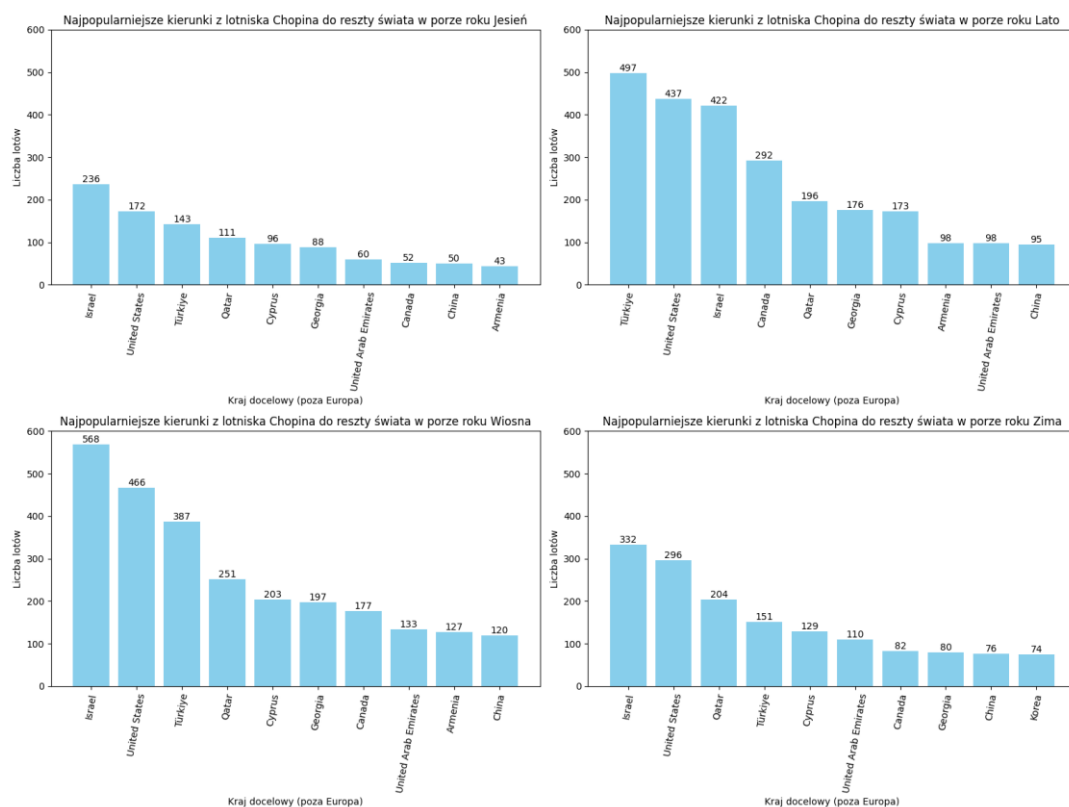
Rysunek 20 Najbardziej oblegane państwa, z których podróżuje się do lotniska Chopina w zależności w ciągu roku



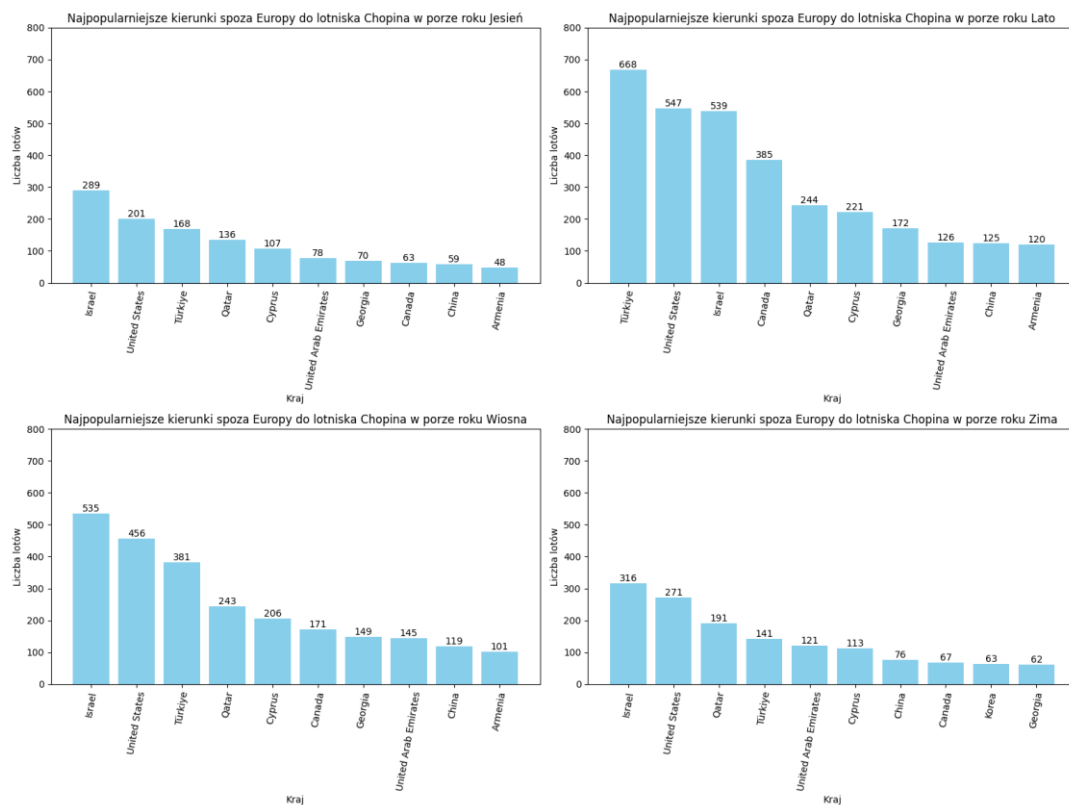
Rysunek 21 Najpopularniejsze kierunki państw z Europy w zależności od pory roku



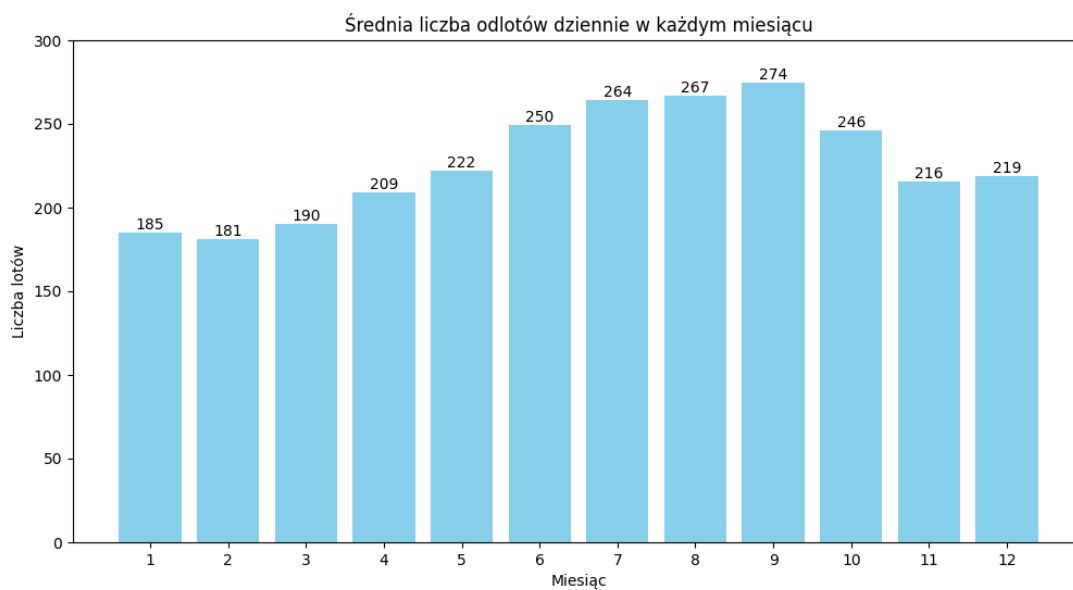
Rysunek 22 Najbardziej oblegane państwa z Europy, z których podróżuje się do lotniska Chopina w zależności od pory roku



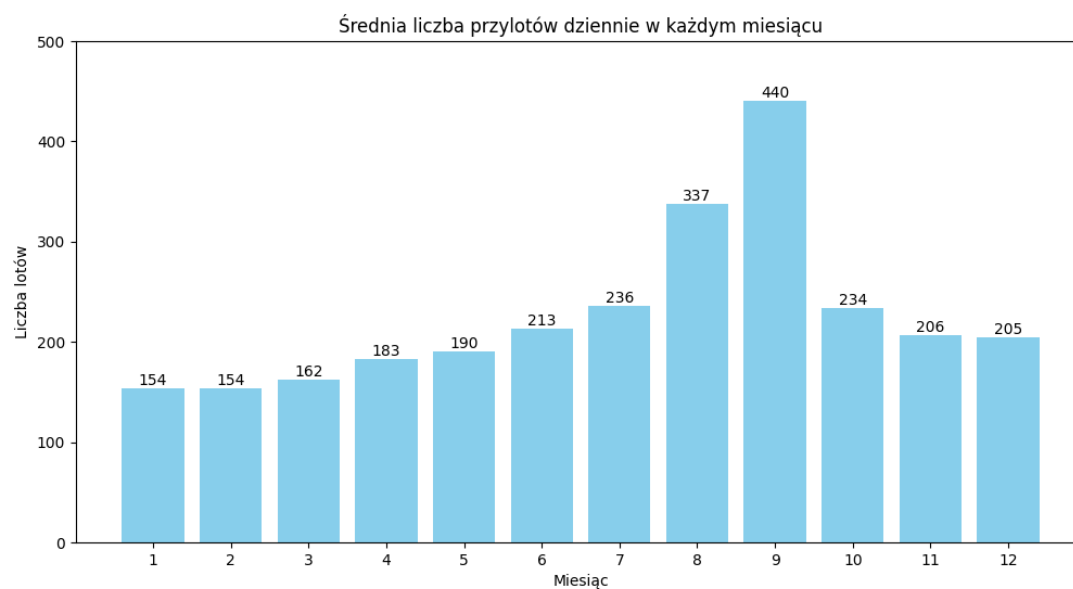
Rysunek 23 Najpopularniejsze kierunki państw spoza Europy w zależności od pory roku



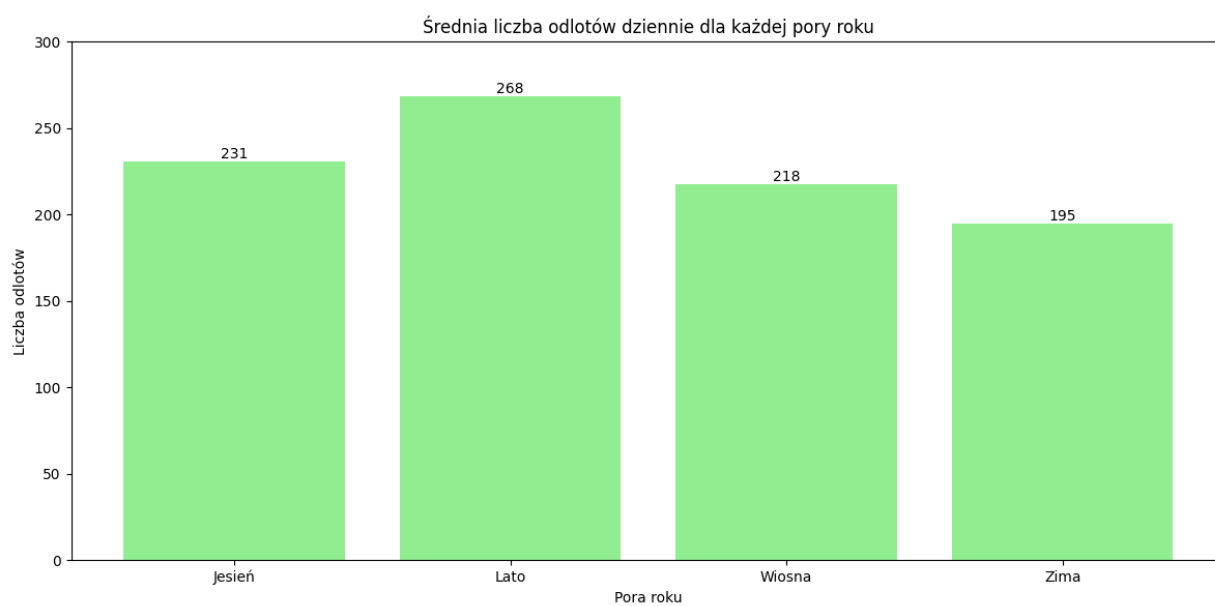
Rysunek 24 Najbardziej oblegane państwa spoza Europy, z których podróżuje się do lotniska Chopina w zależności od pory roku



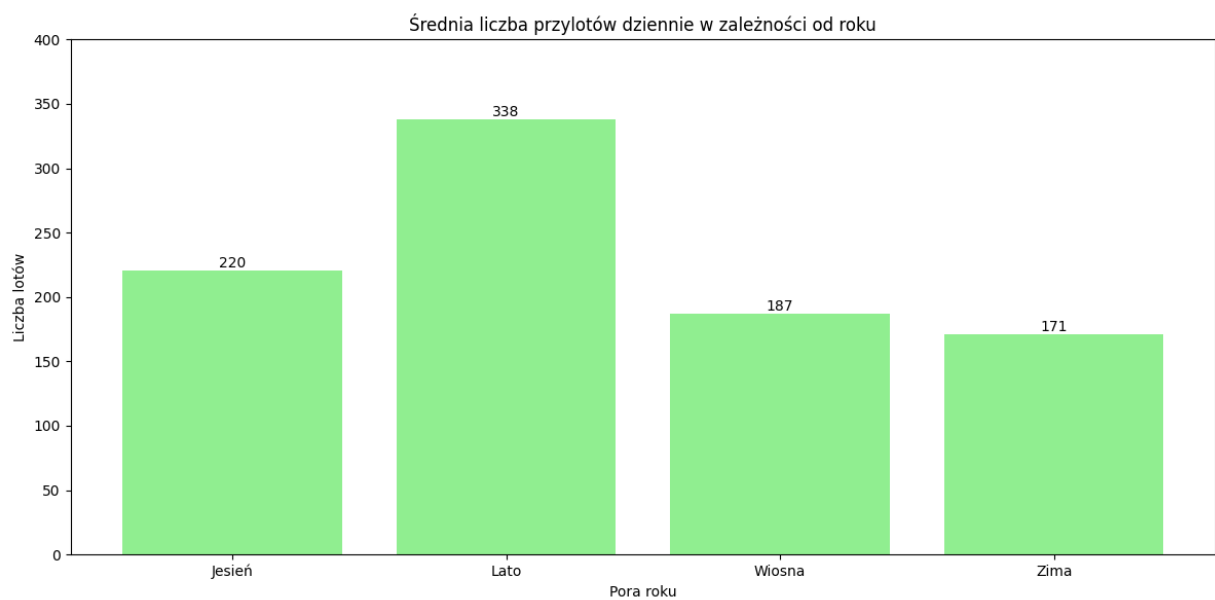
Rysunek 25 Analiza wykonywanych operacji lotniczych odlotów z lotniska w zależności od miesiąca



Rysunek 26 Analiza wykonywanych operacji lotniczych przylotów na lotnisko w zależności od miesiąca



Rysunek 27 Analiza wykonywanych operacji lotniczych odlotów z lotniska w zależności od pory roku



Rysunek 28 Analiza wykonywanych operacji lotniczych na lotnisko w zależności od pory roku



## 7. Wnioski

Analiza danych wykazała, że lotnisko Chopina jest najbardziej obciążone w sezonie letnim. Drugim okresem o wysokiej intensywności ruchu lotniczego jest wiosna. Jesień i zima charakteryzują się względnie spokojnym ruchem w porównaniu do lata i wiosny.

(Rysunek 13, Rysunek 14)

Jesień	Lato	Wiosna	Zima
25 613	55 667	49 353	33 014

*Tabela 1 Ilość operacji lotniczych w zależności od sezonu.*

Pora roku	Jesień	Lato	Wiosna	Zima
Odloty	12163	24691	26546	17580
Przyloty	13450	30976	22807	15434

*Tabela 2 Ilość operacji z podziałem na przyloty i odloty*

Dalsza analiza wykazała, że z lotniska Chopina najczęściej podróżuje się do Europy w sezonie wiosennym. Na drugim miejscu plasuje się Azja, jednak różnica między liczbą operacji lotniczych do Europy a Azji wynosi około 20 tysięcy. Przyloty do lotniska Chopina najczęściej odbywają się w okresie letnim, zarówno z Europy, jak i z Azji.

(Rysunek 15, Rysunek 16)

Podróże z lotniska Chopina najchętniej odbywają się do krajów Europy Wschodniej, a następnie Europy Zachodniej. Najczęściej wybierane kierunki poza Europą to Zachodnia Azja oraz Ameryka Północna. Podobna sytuacja występuje w przypadku przylotów do Chopina.

(Rysunek 17, Rysunek 18)

Najpopularniejsze państwa, do których podróżuje się z lotniska Chopina, to przede wszystkim loty wewnątrz krajowe, a następnie Niemcy, Wielka Brytania, Francja, Włochy, Holandia i Szwajcaria. Z kolei przyloty do Chopina najczęściej odbywają się z Niemiec, a podróże wewnątrz krajowe są na drugim miejscu. Niezależnie od pory roku dominują podróże wewnątrz krajowe w przypadku lotów do Europy, a następnie do Niemiec. W przypadku przylotów, w każdej porze roku dominują Niemcy, z wyjątkiem lata, kiedy Polska jest na pierwszym miejscu. (Rysunek 19, Rysunek 20, Rysunek 21, Rysunek 22)

Poza Europą, najczęstsze operacje lotnicze w zależności od pory roku dotyczą Stanów Zjednoczonych, Turcji, Izraela, Kataru, Cypru oraz Kanady. Podobna sytuacja dotyczy przylotów.

(Rysunek 23, Rysunek 24)

W 2018 roku można zaobserwować zmiany w liczbie wykonywanych lotów. Tendencja jest wzrostowa aż do października, z wyraźną dominacją okresu letniego, kiedy to najwięcej osób podróżuje w celach wypoczynkowych. Po wrześniu widoczny jest spadek liczby lotów, co związane jest z rozpoczęciem roku szkolnego. W tym czasie wzrasta prawdopodobnie liczba lotów biznesowych, aż do lipca następnego roku.

(Rysunek 25, Rysunek 26, Rysunek 27, Rysunek 28)

## 8. Perspektywy dalszego rozwoju projektu

Projekt posiada liczne perspektywy rozwoju. Pierwszą z nich jest rozszerzenie analizy na wszystkie lotniska w Polsce, co umożliwiłoby szczegółową analizę obciążenia poszczególnych portów lotniczych. Kolejnym krokiem mogłoby być zbieranie danych z kolejnych lat dla wszystkich lotnisk, co pozwoliłoby na analizę zmian w liczbie lotów oraz ocenę potrzeby stworzenia centralnego portu lotniczego (HUB). Wykorzystanie technik uczenia maszynowego pozwoliłoby na prognozowanie przyszłych trendów lotniczych i identyfikację potencjalnych krajów, z którymi centralny port lotniczy mógłby nawiązać połączenia.

Największą przeszkodą jest niestety czasochłonny proces zbierania danych.

## 9. Spis ilustracji i tabel

Rysunek 1 Przegląd danych (dla typy position) .....	4
Rysunek 2 Wycinek kodu, kryteria wyszukiwania połączeń z lotniska Chopina .....	4
Rysunek 3 Wycinek kodu, kryteria wyszukiwania połączeń do lotniska Chopina .....	4
Rysunek 4 Czas wyszukiwania danych dla 1 stycznia 2018 roku.....	5
Rysunek 5 Dane zawierające połączenia do lotniska Chopina .....	5
Rysunek 6 Dane zawierające połączenia z lotniska Chopina.....	5
Rysunek 7 Problematiczne dane.....	5
Rysunek 8 Użyte biblioteki .....	6
Rysunek 9 Wczytanie danych.....	6
Rysunek 10 Mapowanie danych.....	6
Rysunek 11 Sprawdzenie poprawności mapowania.....	7
Rysunek 12 Podział miesięcy na pory roku .....	7
Rysunek 13 Liczba odlotów z lotniska Chopina w ciągu roku .....	9
Rysunek 14 Liczba przylotów do lotniska Chopina w ciągu roku .....	9
Rysunek 15 Liczba lotów do kontynentów w zależności od pory roku .....	10
Rysunek 16 Liczba przylotów z kontynentów do lotniska Chopina w zależności od pory roku .....	10
Rysunek 17 Najpopularniejsze kierunki w zależności od pory roku .....	11
Rysunek 18 Najpopularniejsze regiony świata, z których podróżuje się do lotniska Chopina w zależności od pory roku.....	11
Rysunek 19 Liczba lotów do dziesięciu najpopularniejszych państw.....	12
Rysunek 20 Najbardziej oblegane państwa, z których podróżuje się do lotniska Chopina w zależności w ciągu roku .....	12
Rysunek 21 Najpopularniejsze kierunki państw z Europy w zależności od pory roku.....	13
Rysunek 22 Najbardziej oblegane państwa z Europy, z których podróżuje się do lotniska Chopina w zależności od pory roku .....	13
Rysunek 23 Najpopularniejsze kierunki państw spoza Europy w zależności od pory roku ....	14
Rysunek 24 Najbardziej oblegane państwa spoza Europy, z których podróżuje się do lotniska Chopina w zależności od pory roku .....	14
Rysunek 25 Analiza wykonywanych operacji lotniczych odlotów z lotniska w zależności od miesiąca.....	15
Rysunek 26 Analiza wykonywanych operacji lotniczych przylotów na lotnisko w zależności od miesiąca.....	15
Rysunek 27 Analiza wykonywanych operacji lotniczych odlotów z lotniska w zależności od pory roku .....	16
Rysunek 28 Analiza wykonywanych operacji lotniczych na lotnisko w zależności od pory roku.....	16
 Tabela 1 Ilość operacji lotniczych w zależności od sezonu. ....	17
Tabela 2 Ilość operacji z podziałem na przyloty i odloty .....	17

Link do Github: [Adrint/ONS \(github.com\)](https://github.com/Adrint/ONS)