
LOVELY PROFESSIONAL UNIVERSITY

**Name : Tolupunoori Rajshekar.**

**Section:K21UWB54**

**Course-Name :EDA Project**

**Course Code - CSE353**

# Continuous

# Assessment - 1

# 1.Domain and Topic Knowledge:

My dataset is about all the airports in the world.

We get knowledge about all the airports in the world. We can see whether the airport is operational or not and the exact coordinates of the airport.

This dataset contains information about airports globally.

It can be used to analyze the geographic distribution and density of airports around the world and in different countries/regions.

Analysis can be done on airport infrastructure features like runway size, airport facilities, and elevations.

Airports can be categorized by size, usage, and whether they have regular airline service.

Network analysis can identify connectivity between airports based on routes.

Airport locations can be linked with demographic, economic, and transportation data for analysis.

Usage includes understanding geographic accessibility to aviation infrastructure, both between and within countries.

Can support infrastructure planning, route development, and strategic decision making for airlines and airports.

Enables a wide range of geographical, infrastructure, operational, and network analyses of the global airport system.

Includes data on airport locations, infrastructure, operational characteristics, and identifiers.

Geographic coordinates provide exact latitude and longitude for mapping airports.

Airport classifications indicate size and usage like airliner, general aviation, etc.

Country and region codes link airports to specific geographical locations.

Runway data gives insight into airport capabilities and limitations.

Elevations provide information relevant to approach and landing.

Airline service field shows if regular passenger service is offered.

Allows analysis of airport distributions, connectivity, infrastructure, and capabilities at global and regional levels.

Can be integrated with demographic, economic, transportation, and other data for more advanced analytics.

Supports geographical visualization, network analysis, infrastructure planning, and operational studies of the worldwide airport system.

# 2.Data Understanding

1. **id** - Unique integer identifier for each airport. Useful as the primary key.

2. **airport_ident** - Alphanumeric code that uniquely identifies each airport globally. Standardized across aviation databases.

3. **type** - Categorizes airport by size and usage - small_airport, heliport, seaplane_base etc. Smaller airports have fewer/shorter runways.

4. **name** - Official textual name of the airport. Allows joining additional airport data.

5. **latitude_deg** - Latitude coordinate in decimal degrees specifies exact geospatial location. Enables mapping.

6. **longitude_deg** - Longitude coordinate in decimal degrees specifies exact geospatial location. Enables mapping.

7. **elevation_ft** - Elevation in feet above mean sea level. Relevant for takeoff requirements.

8. **iso_country** - Standard 2-letter country code based on ISO standard. Group airports by country.

9. **iso_region** - Standard code for state/region within a country. Provides grouping within the country.

10. **municipality** - City/town where the airport is located. Helps associate demographic and economic data.

11. **scheduled_service** - Indicates if the airport offers regular scheduled passenger service. Categorizes usage.

12. **gps_code** - Unique GPS code assigned for navigation systems. Alternative standardized identifier.

13. **local_code** - Local IATA 3-letter code used in the area. Simplified alternative identifier.

14. **description** - Text indicating communication protocol - CTAF, UNICOM etc. Specifies how pilots communicate.

15. **frequency_mhz** - Numeric radio frequency for communications. Standard aviation radio frequencies used.

16. **runway_length_ft** - Length of longest runway in feet. Indicates airport capabilities and limitations.

17. **runway_width_ft** - Width of longest runway in feet. Also indicates capabilities.

18. **runway_surface** - Surface type - asphalt, gravel etc. Describes runway composition.

19. **runway_lighted** - Indicates if the runway is lighted for nighttime use. Safety and operational data point.

20. **runway_closed** - Specifies if the longest runway is closed. Indicates operational status.

# 3.Reasons For Choosing This Data

It provides a comprehensive list of airports globally, enabling analysis of the worldwide aviation infrastructure and networks.

The inclusion of airport location coordinates allows geographical mapping and analysis of airport distribution patterns.

Details on airport characteristics like runways, lighting and communications enable analysis of infrastructure capabilities and limitations.

Classification by airport type and size facilitates analysis of airport roles and connectivity for things like regional access.

Information on scheduled airline service helps categorize airports by passenger traffic and commercial operations.

The standardized codes provide consistent identifiers that can be used to join additional datasets for more detailed analysis.

Elevation data is relevant for studies on operational requirements and obstacles for approaches and takeoffs.

The data can be used for visualizations, network analysis, infrastructure planning, identifying underserved regions, and more.

It is a freely available dataset that provides a rich, multifaceted view of global airports for both academic study and practical use cases.

# 4.Questions For Analysis

1. What is the unique numeric ID for the airport with the longest runway?
2. What is the alphanumeric code identifying the airport with the highest elevation?
3. What is the character name of the airport with the most scheduled airline service?
4. What is the latitude coordinate of the airport in the United States with the lowest elevation?
5. What is the longitude coordinate of the airport in Europe with the longest runway?
6. What is the country code for the airport with the most heliports?
7. What is the region code for the airport with the largest municipal area?
8. What is the character name of the city/town where the airport with the most scheduled airline service is located?
9. Does the airport with the highest elevation have regular scheduled airline service?
10. What is the alphanumeric code used by GPS systems to identify the airport with the most heliports?
11. What is the local IATA code used in the local area to identify the airport with the longest runway?
12. What is the character descriptor for the airport communication frequencies at the airport with the most scheduled airline service?
13. What is the radio frequency for communications at the airport with the highest elevation?
14. What is the length of the longest runway at the airport with the most heliports?
15. What is the width of the longest runway at the airport with the lowest elevation?
16. What is the surface type of the longest runway at the airport with

the most scheduled airline service?

17. Is the longest runway at the airport with the highest elevation lighted for night operations?
18. Is the longest runway at the airport with the most heliports closed?
19. What is the average elevation of all airports in the United States?
20. What is the average number of scheduled airline services for all airports in Europe?
21. What is the most common airport type in Asia?
22. What is the least common airport type in Africa?
23. What is the median runway length for all airports in South America?
24. What is the interquartile range of runway widths for all airports in North America?
25. What is the correlation between airport elevation and runway length?