

AIRDASHER

**ANALYTICAL INTELLIGENCE FOR ROUTE DEVELOPMENT
AND STRATEGIC HANDLING OF ECONOMIC RESOURCES**

DETAILED REPORT

**SUBMISSION FOR R² DATA LABS RESPONSIBLE GROWTH
HACKATHON**

TEAM: π-thon (SAKSHAT RAO)

PROBLEM STATEMENT:

Economies & industries are starting to spread across the entire Indian subcontinent and are growing at remarkable speed. This has caused increasing demands for several different travel routes between not only major tier-I cities like New Delhi, Mumbai & Bengaluru but also tier-II cities.

Not only this, but rising incomes paired with competitive airport fares have started to bridge the gap between unaffordable & affordable air travel options, especially for the middle-class. Additionally, the post-pandemic world has given airlines & airports an opportunity to not only revive but also grow their business for the next few years.

Hence, air travel in India is seeing good growth in terms of air traffic and it is justified for airlines & airports to capitalize on this growth by fulfilling the growing demand. The best way to do this is to open new routes and thereby new income avenues.

Opening new routes is a business decision which airlines, along with the backing & support of airports & tourism boards, should proactively undertake regularly. But in these times of growth, it is too big an opportunity cost to not expand an airline's network to include high-growth destinations. However, opening new routes is a capital-heavy and financially-risky decision.

As per this [article](#), success rates for low-cost US carriers was 50-70%, meaning 30-50% of newly developed routes did not remain operational beyond two years of inauguration. This indicates scope to further improve the route development process and not rely on wishful thinking, but rather on market reality.

Another area of concern is that new route development requires data which accurately represents customer demands & desires. Data should accurately indicate which routes will be well-received and which ones are not really required. Direct customer feedback is probably the best kind of data for this. This kind of data can not only be collected through surveys but also through flight searches and unfulfilled searches. This kind of data can help identify which routes seem to be required by customers but are not available in the market.

However, direct customer feedback is difficult, time-consuming and expensive to collect. Perhaps macro-economic factors of a city can also be used for forecasting route demands. Though probably not as advantageous as direct customer feedback,

macro-economic factors can be collected through various data sources easily and can help in gauging long-term forecasts.

Finally, route development does not only consider route demand from the customer side but also feasibility & profitability from the airline's side. A route which is expected to see high demand in future is of no use until airlines can translate the excessive demand into profits. And the translation of demands into profit requires a highly parameterized approach which looks at various aspects of the airline's business model. This must also be an essential part of a route development program.

RELEVANCE TO THE THEME:

A problem statement expecting a route development program which looks at macro-economic factors and a parameterized approach is highly relevant to the theme of "Identification of new air travel routes in emerging markets". Such a problem encourages a pro-active agile mindset, helping in aggregating market & competitor insights and building a data-driven business.

Further, this problem is directly involved with the civil aerospace sector where Rolls-Royce is one of the biggest players in terms of building power & propulsion solutions for safety-critical application while also focusing on electric aviation and other sustainable solutions.

AIRDASHER:

Now would be a good time to introduce AIRDASHER, my proposal to solve the above problem statements. AIRDASHER stands for -

"Analytical Intelligence for Route Development And Strategic Handling of Economic Resources"

Following are the key terms of my solution -

- Use of advanced **Analytics** & visualizations along with Artificial **Intelligence** from Machine Learning models
- **Route Plan Development** using data-driven evaluation of current & future demand
- Experimentation with different **strategies** to enter and compete in the market

- Effective **handling** of airline's **economic resources** like capital & fleet

MARKET:

The following table concisely describes the market that AIRDASHER will be targeted towards -

MARKET	MARKET SIZE	MARKET GROWTH
Indian Airlines	8 Major + 6 Scheduled Commuter Airlines with total of 700+ operating fleet size	India is the 3rd largest & fastest growing aviation sector in the world. Among 14 main Indian airlines, 1400+ new orders have been made to increase total fleet by ~200%.
Route Development Analytics	Global Aviation Analytics market worth ~\$2 billion dollars while Route Profitability Software market worth ~\$10 billion	Global Aviation Analytics market expected to grow to ~\$5 billion by 2028 at CAGR of 12% while Route Profitability Software market expected to grow to \$27 billion by 2030 at CAGR of 10% .

Airlines are actively looking to expand their networks, as is evident by the recently-placed large orders of planes by **Air India & IndiGo**. The global aviation analytics market is also an up-and-coming market which will be in high-demand once airlines actively look at expanding networks.

COMPETITIONS:

AIRDASHER aims to provide following applications -

- New Route Development: To Airlines
Exploring new routes and gauging feasibility/profitability
- New Route Development: To Airports
Exploring new routes and supporting/backing airlines to operationalize this route
- New Route/Airport Development: To regional government/tourism boards
Identifying new airport candidates to not only fulfill air traffic demand but also stimulate new demand

- **Aviation Analytics:** To aircraft lessors & manufacturers/OEMs
To track how different types of aircrafts are being used by airlines in different regions of the world and accordingly use this data in their decision-making

The following table concisely describes the direct & indirect competition for AIRDASHER's applications -

APPLICATIONS	TARGETED MARKET	COMPETITION
Route Opportunity Analysis	Airlines	Jeppesen Aircraft Routing, ASM Passenger Route Forecast
Route Cost Analysis & Optimizations	Airlines	ASM Passenger Route Forecast Tool
New Route/Airport Development	Airports	Avia Solutions' Route Development, ASM Airport Long Term Passenger Forecast
New Route/Airport Development	Regional Governments / Tourism Boards	Avia Solutions' Route Development
Aviation Analytics	Aircraft Lessors & Manufacturers/OEMs	Cirium Airline Routes

The main competitor would probably be ASM Global Route Development tool. This tool can analyze market potential of a selected route and help in determining potential performance of a route, frequency, timings & aircraft type. Additionally, it can provide forecasts on passenger demand, capacity, movements, etc for an airport with forecasts being supported by non-commercial & cargo projections. These forecast segmentations include long-haul, short-haul, domestic carrier, regional carrier & low-cost carrier.

Other competitors provide indirect or partial competition. For example, Cirium provides satellite-based flight tracking data, Rezcomm provides an interface for collecting customer feedback at airports, Jeppesen provides assistance in operational aspects of a flight while Avia provides airports & regional partners with assistance in developing and proposing new routes to airlines.

Overall, there are only a handful of direct competitors in the market. The main distinction from the competition would need to be a complete end-to-end solution, the ability to use macro-economic factors to make long-term forecasts more insightful &

accurate while also being able to provide a highly parameterized model for cost analysis. This would enable any airline to get customized results based on the parameters they input, which would depend on the airline's business strategy and goals.

PROPOSED TECHNICAL SOLUTION:

Now, going slightly into the more technical aspects of how AIRDASHER achieves its objectives and applications, the solution can be divided into following parts -

1. New Route Development
 - a. City Selection
 - b. Route Selection
 - c. Cost & Resource Analysis
2. New Airport Development
3. Civil Aviation Statistics
4. Web Server

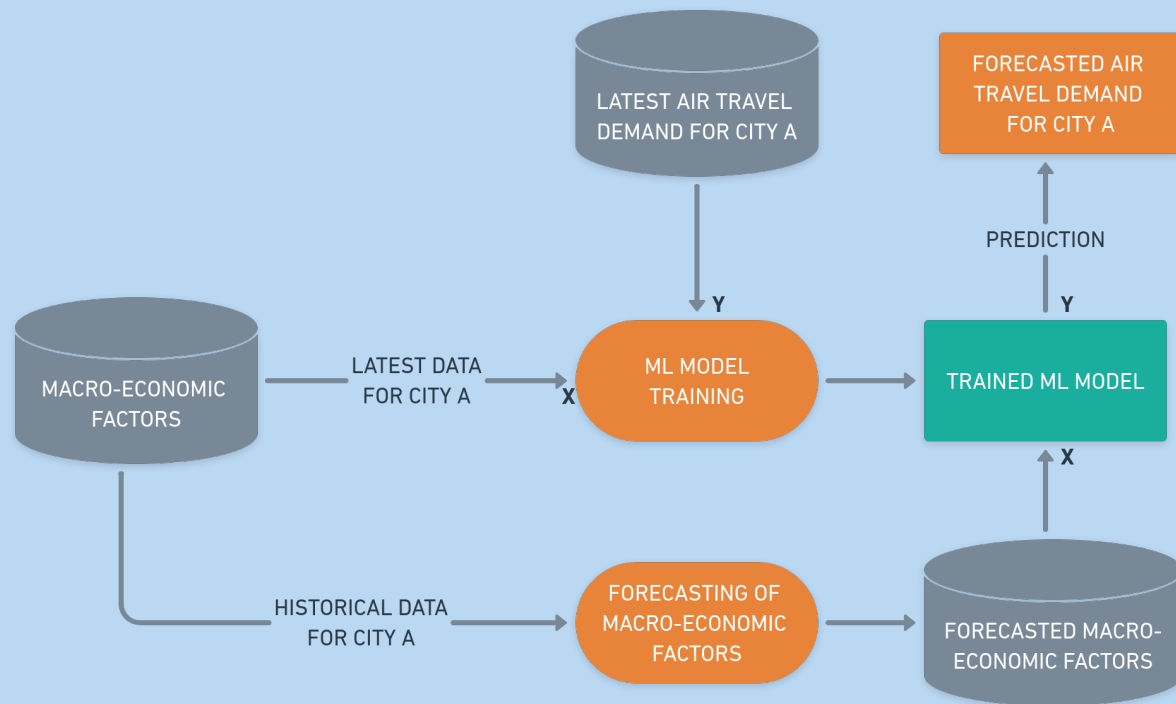
These applications are accessible to users via a website which has been hosted online at <http://airdasher.pythonanywhere.com>. This website is hosted on PythonAnywhere and it is highly suggested to try the website and its different features to appreciate how this solution could help airlines with their route development process.

Let us discuss the individual technical parts of AIRDASHER -

New Route Development:

City Selection:

The below diagram shows the overall flow of how AIRDASHER is able to forecast air travel demands for different cities -



We use four different categories of macro-economic factors - **Population, Education, Economics & Tourism.**

The datasets used here are -

Dataset Name	Reference	Metadata	Information
City-pair wise Domestic Data	https://www.dgca.gov.in/digigov-portal/?page=monthlyStatistics/259/4751/html&main259/4184/servicename	Year: 2022 Source: Directorate General of Civil Aviation	Provides domestic passenger & cargo data for different city-pairs
Indian Districts Economic Data	https://censusindia.gov.in/nada/index.php/catalog/42526/download/46152/A-1_NO_OF_VILLAGES_TOWNS_HOUSEHOLDS_POPULATION_AND_AREA.xlsx	Source: 2011 Census, Office of the Registrar General & Census Commissioner, India Ministry of Home Affairs, Government of	Provides GDP information & history for different Indian districts

		India	
Population/Household/Area Data	https://censusindia.gov.in/nada/index.php/catalog/42526/download/46152/A-1_NO_OF_VILLAGES_TOWNS_HOUSEHOLDS_POPULATION_AND_AREA.xlsx	Source: 2011 Census, Office of the Registrar General & Census Commissioner, India Ministry of Home Affairs, Government of India	Provides population, towns, households & area information for different Indian districts
Decadal Variation in Population	https://censusindia.gov.in/census.website/data/census-tables# (table PC11_A02: A-02: Decadal variation in population 1901-2011)	Source: 1901 to 2011 Census, Office of the Registrar General & Census Commissioner, India Ministry of Home Affairs, Government of India	Provides population history for different Indian districts
Latest Indian City Population Data	https://worldpopulationreview.com/countries/cities/india	Year: 2023, source: https://worldpopulationreview.com/	Provides latest population data for Indian cities
Education Data	https://censusindia.gov.in/census.website/data/census-tables# (tables 1991-C02T: C-2 Age Sex And Educational Level - All Areas, PC01_C08a: C-08 Appendix: Education level graduate and above by sex for population age 15 and above (all) - Appendix, PC11_C08a: C-08 Appendix: Educational level graduate and above by sex for population age 15 and above	Source: 1991, 2001 & 2011 Census, Office of the Registrar General & Census Commissioner, India Ministry of Home Affairs, Government of India	Provides number & history of graduates for different age groups in Indian districts

	(total) - Appendix)		
Tourism Data	<p>Domestic & Foreign visitors to Centrally Protected Ticketed Monuments Data: Data manually collected from</p> <p>https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwiY35yy0f39AhVISWwGHXL6CwIQFnoECBIAQ&url=https%3A%2F%2Ftourism.gov.in%2Fsites%2Fdefault%2Ffiles%2F2022-09%2FIndia%2520Tourism%2520Statistics%25202022%2520%2528English%2529.pdf&usg=AOvVaw1-4KC-GKJllotlsIxFFMe, https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwjTk-rA0f39AhVGT2wGHcHyDEEQFnoECAwQAQ&url=http%3A%2F%2Ftourism.gov.in%2Fsites%2Fdefault%2Ffiles%2F2022-09%2FIndia%2520Tourism%2520Statistics%25202021%2520%25281%2529.pdf&usg=AOvVaw3h3V-3AQKvZHtdbq5QSM7j, https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwjShrDE0f39AhWJSGwGHVQxBooQFnoECBAQAQ&url=https%3A%2F%2Ftourism.gov.in%2Fsites%2Fdefault%2Ffiles%2F2021-05%2FINDIA%2520TOURISM%2520STATISTICS%25202020.pdf&usg=AOvVaw0INamqon9LOB2ASVUCEIb0</p>	<p>Source: India Tourism Statistics Reports 2022, 2021 & 2020 (Chapter: Domestic Tourism, Table: Visitors to Centrally Protected Ticketed Monuments), Ministry of Tourism, Government of India</p>	<p>Provides domestic & international tourist information & history for several centrally protected ticketed monuments in India</p>

In a nutshell, several macro-economic factors for most of the tier-I & tier-II cities are collected and fed to a Machine Learning model as features to predict the total air traffic

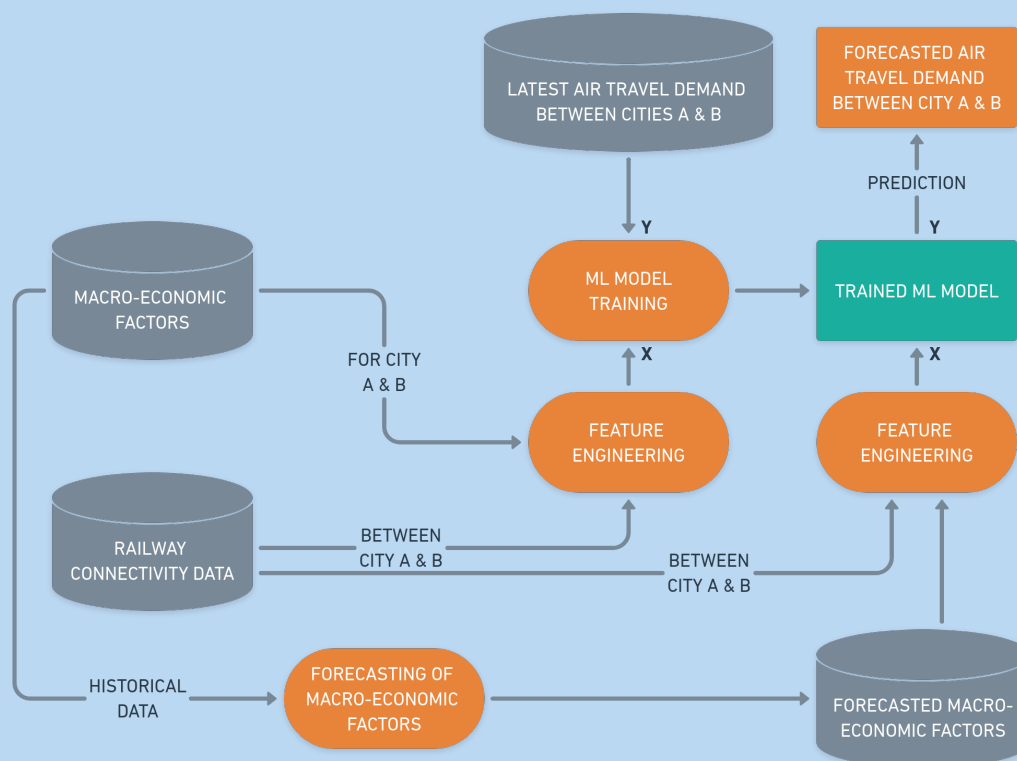
coming in and out of the city. We assume that the model is able to identify the important features, relationships & patterns from the macro-economic data.

For each factor, we then use its historical information to fit an exponential curve. Using this curve, we can forecast the values of these macro-economic factors in the future. Using these forecasted values, we feed it into the trained model to get forecasted air traffic demands for each airport.

AIRDASHER then sorts the cities having highest forecasted demand as well as highest forecasted growth in demand and lists such cities as candidates to add to an airline's network.

Route Selection:

The below diagram shows the overall flow of how AIRDASHER is able to forecast air travel demands for different routes for a given city -



The datasets used here are -

Dataset Name	Reference	Metadata	Information
Flight / Airport Data	Manually created from https://www.flightconnections.com/	<p>Assumptions:</p> <p>(1) All airlines considered (no bias involved)</p> <p>(2) Prices are for one-way economy for 1 adult on May 1, 2023 (chosen date is far enough to avoid near-date fluctuations in prices)</p> <p>(3) Non-stop flights only, 1+ connection flights are not being considered</p>	Info about over 850+ flights & 125+ airports in Indian & APAC regions
Railway Connectivity Data	https://github.com/datameet/railways , https://github.com/geohacker/railways	Year: Unknown (seems to be 2016), source: GitHub (Sanjay Bhangar [Twitter: @sanjaybhangar] & Sajjad Anwar [Twitter: @geohacker])	Provides information about all trains in the Indian Railways
District Centroid Data	https://www.kaggle.com/datasets/sirpunch/indian-census-data-with-geospatial-indexing?select=district+wise+centroids.csv	Source: 2001 & 2011 census (https://sumit-ara.medium.com/plotting-weighted-mean-population-centroids-	Information about the centroids of different Indian districts

		on-a-country-map-22da408c1397)	
Station Code Data	https://en.wikipedia.org/wiki/List_of_railway_stations_in_India	Source: Wikipedia	Information about station codes

In a nutshell, for all city-pairs in the domestic passengers data, we collect macro-economic information for both cities, extract flight duration & distance for a flight between the two cities & collect data about railways connectivity between the two cities. We train another model with all this data as features trying to predict air traffic for the given city-pair route.

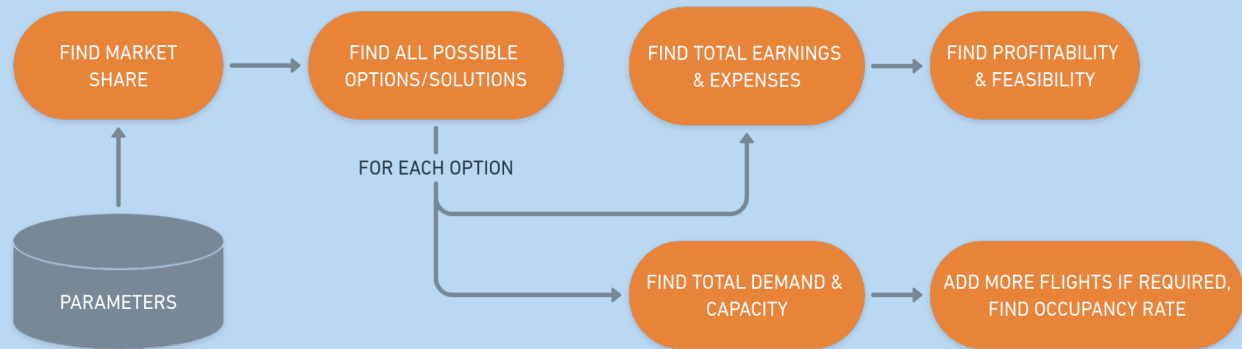
Similar to the last step, we forecast the macro-economic factors for both cities. We also assume railway data to be growing at a certain rate and hence, can forecast for that data as well. These forecasted values are fed into the trained model to get forecasted air traffic demands for the given city-pair.

For AIRDASHER to analyze all routes from the selected city to either the hubs of the selected airline's network or all airports in the selected airline's network (based on a parameter), it collects the macro-economic data, flight duration/distance & railway connectivity information for each route and gets a forecasted air traffic demand for the route.

AIRDASHER then sorts the possible routes having highest forecasted demand as well as highest forecasted growth in demand and lists such routes as candidates to connect the selected city to the selected airline's network.

Cost & Resource Analysis:

The below diagram shows the overall flow of how AIRDASHER is able to perform cost/resource analysis for the selected route -



In a nutshell, AIRDASHER collects the different parameters from the user and checks different combinations of frequency & types of aircrafts. For example, it checks for 2 turboprop flights or 1 narrowbody+1 turboprop flight operating on the given route. For each scenario, it calculates the -

- Market Share (depends on flight fare & number of flights)
- Total Demand (depends on selected route & market share)
- Total Capacity (depends on frequency & number of flights)
- Total Expenses (depends on number & type of flights)
- Total Earnings (depends on flight fare & demand/capacity of flights)

Using this information, it can calculate the profit margin & occupancy rates for the given scenario/option/combination. After iterating through this process for almost all combinations, AIRDASHER sorts the options having highest feasibility or profitability.

New Airport Development:

The process of finding candidates for new airports is very similar to the process of finding candidates for cities to be added to an airline's network. We simply fetch the required macro-economic data for those cities which do not have airports. By feeding this data to our models, we can forecast the air traffic demand for such cities, even though they do not have airports. This is one of the advantages of using Machine Learning intelligence.

For all such tier-I/II cities not having an airport, we can find the theoretical air traffic demand for such cities. AIRDASHER then sorts the cities having highest demand & growth in air traffic demand and lists them as candidates to establish a new airport in.

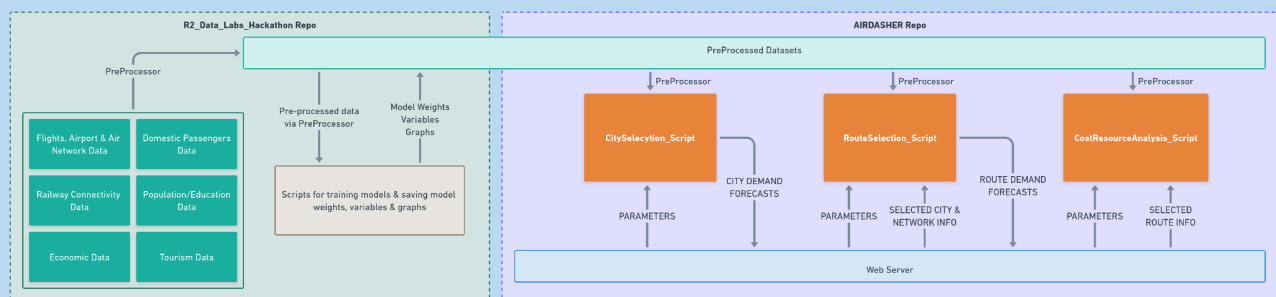
Civil Aviation Statistics:

Using the comprehensive manually-built dataset of 850+ flights connecting 125+ Indian & APAC cities, we can analyze several aspects of the current civil aviation sector. We can derive insights about how different aircraft types are being used, find a model to detect underpricing & overpricing and also compare connectivity towards major tier-I cities & other tier-II cities.

Such insights along with several other can help the targeted markets of aircraft lessors & manufacturers/OEMs understand how aircrafts are being used in different geographies. This in turn can be valuable inputs going into their respective decision-making process.

Web Server:

Below figure shows entire workflow from how data is preprocessed to how web server is able to access the different information -



BUSINESS MODEL:

The business model for AIRDASHER is given as two opportunities of offering AIRDASHER as -

- **End-to-End Route Development Tool**
 - Targeted mainly towards different airlines in Indian/APAC regions
 - Application would include air travel demand analysis/forecasting along with parameterized cost & resource analysis. This will help airlines in route development process end-to-end
- **Civil AviationSector Analytics Tool**
 - Targeted towards Airports, Regional Governments, Tourism Boards, Aircraft Lessors, Manufacturers & OEMs
 - Application would include Air travel demand analysis & forecasting

Estimated Opex:

Given below is a table listing the possibly required OpEx costs for operating & maintaining the AIRDASHER website -

BUSINESS FEATURE	OPEX
WEB SERVER (RENT), DOMAIN NAME, SSL CERTIFICATES, WEB DESIGN	\$100 / MONTH
MARKETING & SEO	\$5000 / MONTH
BUSINESS CLOUD STORAGE	\$25 / MONTH

Timeline:

The timeline for bringing AIRDASHER to market could look like this -

- Minimum Viable Product - Would take 3 months
- Product Offering - Would take 9 months

KEY SUCCESS FACTORS:

- It would be a sensible strategy to target a quick entry into the market. Given that the airline industry is slowly reviving and growing beyond its pre-pandemic levels, the need for extremely efficient Route Development tools is also growing and AIRDASHER has the potential to capture a significant portion of the need.
- Using only macro-economic factors would be an indirect way of estimating future customer demands. However, direct customer feedback provides a much better advantage in this aspect. It would seem to be the best trade-off where direct customer feedback & macro-economic data are used together for forecasting. Direct customer feedback can help in accurate short-term forecasts while macro-economic data can help in capturing long-term vision of Indian air travel. Customer feedback could be collected from surveys or could be monitored through flight searches & ad engagement.

- Finally, expert opinion for such a business is of paramount importance given that this business deals directly with the operation, costs, strategies & risks of an airline. Having expert knowledge for the targeted use-cases will help in dealing with the market reality, understanding customer requirements & avoiding potential problems in the future.

CONCLUSION:

Thank you for the opportunity of creating this proposal for the R² Data Labs Responsible Growth Hackathon.

The working website is present at <http://airdasher.pythonanywhere.com>. It is highly suggested to check the website and its applications to better understand how a possible solution like AIRDASHER could help airlines with their New Route Development process.

Given below are few of the other resources which are part of my solution -

- Demo Video - https://youtu.be/_wfrsszOh0
- GitHub repos - <https://github.com/SakshatRao/AIRDASHER> & https://github.com/SakshatRao/R2_Data_Labs_Hackathon

Future Works:

- Use of better regressors than Linear Regressor (Linear Regression is currently used due to memory size issues on PythonAnywhere web server)
- Advanced feature engineering & feature extraction
- Better training of models like hyperparameter tuning
- Use of other types of datasets
- Make use of valuable 2021 census data (delayed to take place in 2023/2024)
- Considering seasonality to maybe get monthly/weekly forecasts
- Make use of direct customer feedback to make more accurate forecasts for passenger demand
- Include model to suggest best timings for new flights (considering connecting demand as well)
- Changing market share in future

Tech Stack:

- Python
 - Numpy
 - Pandas
 - Sklearn
 - Matplotlib
 - Seaborn
- Django
- Plotly
- HTML
- CSS
- Javascript
 - Leaflet

References:

- https://en.wikipedia.org/wiki/Indian_Railways_coaching_stock & <https://www.quora.com/What-is-the-capacity-of-normal-Indian-passenger-train> for estimating number & capacity of different types of coaches on a train on average
- https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwiQp_S3qNP9AhXlwTgGHYuaB5EQFnoECA8QAQ&url=https%3A%2F%2Findianrailways.gov.in%2Frailwayboard%2Fuploads%2Fdirector%2Fstat_econ%2FAnnual-Reports-2020-2021%2FAnnual-Report-English.pdf&usq=A0vVaw2YMsSpEfqTOjBD13N-ZsJl for estimating growth of number of railway coaches, number of trains & running track
- https://en.wikipedia.org/wiki/List_of_airlines_of_India & https://en.wikipedia.org/wiki/List_of_airports_in_India for market size of airlines & airports
- <https://tourism.gov.in/about-us/indian-tourism-offices> for tourism offices info
- <https://www.investopedia.com/ask/answers/050415/what-companies-are-major-players-airline-supply-business.asp> for aircraft manufacturers info
- <https://www.fortunebusinessinsights.com/aviation-analytics-market-102101>, <https://www.alliedmarketresearch.com/aviation-analytics-market>, <https://www.marketsandmarkets.com/Market-Reports/aviation-analytics-market-26274177.html> for market growth
- <https://www.mordorintelligence.com/industry-reports/airline-route-profitability-software-market>, <https://www.marketsandmarkets.com/Market-Reports/aviation-analytics-market-26274177.html>

[26274177.html](#),

<https://www.mordorintelligence.com/industry-reports/airline-route-profitability-software-market> &

<https://www.technavio.com/report/airline-route-profitability-software-market-industry-analysis> for market analysis

- <https://www.soscanhelp.com/blog/how-much-does-business-cloud-storage-cost>, <https://www.hostinger.in/tutorials/website-maintenance-cost> & <https://www.servermania.com/kb/articles/how-much-does-a-server-cost-for-a-small-business/> for business cost estimates
- <https://www.flaticon.com/>, <https://thenounproject.com/> & <https://www.pexels.com/> for icons & pictures
- ChatGPT (<https://openai.com/blog/chatgpt>) for providing city descriptions