STARTUP ECOSYSTEM EXPLORER

**A PROJECT REPORT**

***Submitted by***

PERUMAL L [927623BAD074] SADHANA G [927623BAD095] VENKAT RAGAV N [927623BAD123]

### in partial fulfilment for the award of the degree of

BACHELOR OF TECHNOLOGY

### in

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

**M.KUMARASAMY COLLEGE OF ENGINEERING,**

**KARUR**

**ANNA UNIVERSITY: CHENNAI 600 025**

APRIL 2025

# M.KUMARASAMY COLLEGE OF ENGINEERING

**(Autonomous Institution affiliated to Anna University, Chennai)**

# KARUR – 639 113

BONAFIDE CERTIFICATE

Certified that this project report “STARTUP ECOSYSTEM EXPLORER” is the bonafide work of “PERUMAL L (927623BAD074), SADHANA G (927623BAD095), VENKAT RAGAV N (927623BAD123)” who carried out the project work during the academic year 2024-25 under my supervision. To the best of my knowledge, the work reported herein does not form part of any other project or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

|  |  |
| --- | --- |
| **SIGNATURE**  Ms. S. SIVASANKARI, **SUPERVISOR, CORPORATE TRAINER,**  International Business Machines (IBM). | **SIGNATURE**  Dr. A. SELVI, M.E, Ph.D.,  **HEAD OF THE DEPARTMENT,**  Department of Artificial Intelligence ,  M. Kumarasamy College of Engineering, Thalavapalayam, Karur-639113. |

## TABLE OF CONTENTS

|  |  |  |
| --- | --- | --- |
| **S. No.** | **Section Title** | **Description** |
| **1** | **Introduction** | This project explores startup funding trends using IBM Cognos Analytics. It aims to reveal insights across funding rounds, industries, and regions through interactive visualizations. |
| **2** | **Objectives** | To analyze and visualize startup funding patterns across rounds, industries, and regions using IBM Cognos Analytics. |
| **3** | **Dataset Description** | The dataset contains funding details of 1000 startups, including round type, amount, industry, location, and date. |
| **4** | **Tools and Technologies Used** | Used IBM Cognos Analytics and Excel with a structured data preparation and visualization methodology to analyze startup funding trends. |
| **5** | **Methodology** | Followed a data analytics workflow involving data cleaning, modeling, and visualization using IBM Cognos to uncover funding insights. |
| **6** | **Implementation Steps** | Imported and cleaned the dataset in Excel, created calculated fields, and built interactive dashboards in IBM Cognos using charts like donut, line, heatmap, and bubble map. |
| **7** | **Results and Findings** | Seed and Series A rounds dominate early-stage funding, led by AI and FinTech, with the Bay Area and Bengaluru as top growing hubs. |
| **8** | **Conclusion** | Dynamic insights reveal key trends in startup funding and growth. |
| **9** | **Future Scope** | Adding real-time sentiment analysis and AI-driven predictions for global startup trends. |
| **10** | **References** | IBM Cognos Analytics documentation for dashboard creation, Industry articles on AI and FinTech investment trends |

1. **Introduction**

In the rapidly evolving startup ecosystem, data analytics enables founders, investors, and stakeholders to make strategic decisions based on trends and investment behavior. IBM Cognos Analytics is a business intelligence and visualization platform ideal for interactive dashboards and data storytelling. This project explores global startup investment trends using visual dashboards in Cognos. Our objective is to uncover patterns in funding rounds, growth across industries, and geographic investment distribution by analyzing a curated dataset of 1000 startup funding entries.

## Objectives

1. Funding Trend Analysis – Analyze historical startup funding rounds to identify growth patterns.
2. Sector Performance Insights – Evaluate key industries like AI and FinTech for investment trends.
3. Geographic Funding Distribution – Map and assess regional startup funding hubs globally.
4. Investor Activity Visualization – Create interactive dashboards showing investor participation and deal sizes.
5. Predictive Investment Modeling – Forecast future funding trends based on past data.
6. Startup Growth Impact Assessment – Study how funding influences startup success and ecosystem development.
7. Data-Driven Investment Recommendations – Provide actionable insights for investors and policymakers.

## Dataset Description

This dataset offers comprehensive insights into startup funding activities, capturing details on investment rounds, funding amounts, investor profiles, and startup sectors across multiple regions. It includes metrics such as deal size, funding stage, industry classification, geographic location, and year-over-year growth. Additionally, it records information on investor types and startup valuations, enabling analysis of investment trends and ecosystem dynamics. The dataset supports predictive modeling for funding forecasts and strategic investment decision-making.

* **Website**: Kaggle
* **Data Format**: CSV
* **Total Records**: 1000
* **Total Columns**: 42

**Columns Used in the Analysis:**

|  |  |
| --- | --- |
| **Column Name** | **Description** |
| **Startup ID** | Unique identifier assigned to each startup |
| **Startup Name** | Name of the startup |
| **Industry** | Sector or industry in which the startup operates |
| **Investment Stage** | Stage of investment |
| **Funding Round** | Specific funding round designation |
| **Funding Amount (USD)** | Amount of money raised in US dollars |
| **Investment Year** | Year in which the investment was made |
| **Date of Investment** | |  | | --- | |  |  |  | | --- | | Exact date when the funding was received | |
| **Investor Name** | Name of the individual or organization making the investment |
| **Investor Type** | Category of investor |
| **Country** | Country where the startup is based |
| **Location** | Location of the startup or investment |
| **City** | City where the startup is headquartered |
| **Post-Money Valuation** | Startup’s valuation immediately after the funding round |
| **Employee Count** | Number of employees working in the startup |
| **Founded Year** | |  | | --- | |  |  |  | | --- | | Year the startup was founded | |
| **Startup Age** | Number of years since the startup was founded |
| **Pre-Money Valuation** | Startup’s valuation immediately before the funding round |
| **Revenue (USD)** | Annual revenue of the startup in US dollars |
| **Revenue Growth (%)** | |  | | --- | |  |  |  | | --- | | Percentage growth in revenue over a specified period | |

## Tools and Technologies Used

* **IBM Cognos Analytics** – Utilized for interactive data visualization and dashboard development.
* **Startup Funding Databases** – Sources such as Crunchbase and PitchBook for comprehensive investment data.
* **Data Processing Frameworks** – Tools like Python and Pandas for data cleaning and transformation.
* **Geospatial Mapping Tools** – Used to visualize regional funding distributions.
* **Predictive Analytics Models** – Employed for forecasting funding trends and investment opportunities.

**Microsoft Excel:**

* **Purpose:** Used for initial data exploration, cleaning, and formatting before importing into IBM Cognos Analytics. Excel helped identify missing values, perform quick data summaries, and prepare datasets for detailed analysis.

## Methodology

## Data Collection: Gathering startup funding and investment data from multiple sources like Kaggle, Crunchbase and PitchBook.

## Data Preprocessing: Cleaning, formatting, and structuring the data for analysis using IBM Cognos Analytics.

## Funding Trend Analysis: Identifying investment patterns, key sectors, and geographic funding hubs through data exploration.

## Predictive Modeling: Using historical funding data to forecast future investment trends and startup growth.

## Visualization & Reporting: Creating interactive dashboards to present insights clearly to stakeholders.

## Recommendations: Providing strategic investment and ecosystem development suggestions based on data-driven findings.

## Challenges in Startup Ecosystem Analysis

## Inconsistent or incomplete funding data hinders accurate trend analysis.

## Rapidly evolving industries like AI and FinTech create forecasting difficulties.

## Regional disparities in data availability limit global ecosystem comparisons.

## Identifying reliable early-stage startups is challenging due to limited public information.

## Impact of Startup Analytics on Innovation and Investment

## Data-driven insights support smarter investment decisions and funding allocation.

## Sector-wise trend analysis identifies high-growth industries for targeted support.

## Geographic mapping reveals emerging startup hubs for ecosystem expansion.

## Predictive modeling aids in forecasting startup success and funding opportunities.

## Implementation Steps



## Results and Findings

**Funding Trends and Sector Analysis**

The analysis reveals that Seed and Series A rounds dominate early-stage funding, with AI

and FinTech emerging as the top sectors by average investment size. Geographic hotspots

like the Bay Area and Bengaluru showed strong year-over-year growth, making them

key innovation hubs. Higher investment volumes were often associated with startups

showing rapid revenue growth and large market potential.

**Predictive Insights for Investment Strategy**

Historical funding data enabled forecasting of future investment trends, helping identify

upcoming sectors and regions of interest. Predictive models indicated that early

identification of high-growth startups and sector diversification could enhance

portfolio performance. These insights can support data-driven policymaking and investor

strategies, and contribute to building a robust global startup ecosystem.

## Conclusion

Startup Ecosystem Explorer effectively uncovers key trends in startup funding, highlighting dominant sectors, geographic investment hubs, and growth patterns across various stages. Leveraging IBM Cognos Analytics for visualization and predictive insights, the project empowers data-driven decision-making for investors and policymakers. Future improvements could include integrating real-time funding updates, AI-based success prediction models, and expanded global datasets to further enrich ecosystem analysis.

## Future Scope

* + **Real-Time Funding Data Integration** – Incorporate live updates from funding platforms for timely insights.
  + **AI-Driven Success Prediction** – Use machine learning to forecast startup performance and funding outcomes.
  + **Global Ecosystem Expansion** – Include more countries and regions for broader investment analysis.
  + **Investor Behavior Analysis** – Track and predict investment patterns across sectors and geographies.
  + **Integration with Policy Frameworks** – Support government and institutional decision-making through data-driven insights.

## References

* **Dataset Source – Kaggle: Startup Investment Dataset**  
  <https://www.kaggle.com/datasets>
* **GitHub Repository – Project Code and Visualizations**  
  <https://github.com/your-username/startup-ecosystem-explorer>