



**NAAN
MUDHALVAN
PROJECT BASED
LEARNING**



**MEENAKSHI SUNDARARAJAN ENGINEERING
COLLEGE**

Kodambakkam, Chennai-600024.

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

TOPIC: Estimation Of Business Project

TEAM ID: NM2023TMID07573

FACULTY MENTOR: YAMUNA S

INDUSTRY MENTOR: SHIVAM SHIVHARE

Project submitted by,

ADITHYA MANIKANDAN.J(311520104002),
ASISH AYSHWARIAN.N.R(311520104009),
ELANGO.N(311520104015)
MOHAMED FOWAZ M.A(311520104032)

ABSTRACT

The objective of this data analytics project is to unveil nuanced insights into the financial dynamics of educational institutions, with a specific focus on operating expenses. The project employs a comprehensive approach, dissecting various expenditure categories from 'Annual Payroll' to 'Other Operating Expenses,' providing stakeholders with actionable insights to optimize resource utilization and foster financial sustainability within the education sector. The dataset encompasses a diverse array of expenditure categories, offering a detailed examination of total expenses and their percentage distribution. Critical components such as 'Detailed Employer Costs for Fringe Benefits,' 'Purchased Communication Services,' and 'Depreciation and Amortization Charges' are thoroughly analyzed to provide a holistic understanding of financial patterns within the education sector. Utilizing DB2 metrics and responsive data handling strategies, the project ensures timely access to financial data. Time-based, institutional, and category-specific filters enhance the granularity of the analysis, empowering stakeholders to customize their exploration of operating expenses.

Real-time Data Insights: Implementing mechanisms for real-time data updates and rendering, ensuring stakeholders access the most current financial information.

Query Performance Optimization: Utilizing DB2 metrics to optimize query execution, enhance data retrieval speed, and improve overall system responsiveness.

Interactive Data Exploration: Incorporating user-friendly filters for time, institution, and expense categories, allowing stakeholders to interactively explore and analyze operating expenses.

Customized Insights: Providing flexibility through custom filters, enabling stakeholders to tailor analyses based on specific criteria relevant to their goals.

Cross-Dimensional Analysis: Implementing cross-filtering capabilities to reveal relationships within the data, enhancing the depth of insights derived from the dataset. The data analytics report aims to provide actionable recommendations for stakeholders in the education sector. By understanding the financial dynamics and discerning patterns within each expense category, institutions can make informed decisions, optimize resource utilization, and foster long-term financial sustainability. This project employs advanced data analytics techniques, DB2 metrics, and customizable filters to present a comprehensive analysis of operating expenses in the education sector. The outcomes are intended to guide stakeholders in optimizing financial strategies, ultimately contributing to the overall financial health and sustainability of educational institutions.

Project Report Format

1. INTRODUCTION

- 1.1 Project Overview
- 1.2 Purpose

2. LITERATURE SURVEY

- 2.1 Existing problem
- 2.2 References
- 2.3 Problem Statement Definition

3. IDEATION & PROPOSED SOLUTION

- 3.1 Empathy Map Canvas
- 3.2 Ideation & Brainstorming

4. REQUIREMENT ANALYSIS

- 4.1 Functional requirement
- 4.2 Non-Functional requirements

5. PROJECT DESIGN

- 5.1 Data Flow Diagrams & User Stories
- 5.2 Solution Architecture

6. PROJECT PLANNING & SCHEDULING

- 6.1 Technical Architecture
- 6.2 Sprint Planning & Estimation
- 6.3 Sprint Delivery Schedule

7. CODING & SOLUTIONING

- 7.1 Feature 1
- 7.2 Feature 2
- 7.3 Database Schema (if Applicable)

8. PERFORMANCE TESTING

- 8.1 Performance Metrics

9. RESULTS

- 9.1 Output Screenshots

10. ADVANTAGES & DISADVANTAGES

11. CONCLUSION

12. FUTURE SCOPE

13. APPENDIX

- Source Code
- GitHub & Project Demo Link

1. INTRODUCTION

1.1 Project Overview

The Estimation of Business Project is a comprehensive data analytics initiative designed to provide intricate insights into the financial dynamics of educational institutions. This project centers on the meticulous examination of operating expenses within the education sector, ranging from annual payroll to purchased professional and technical services. It encompasses not only the total expenses but also their percentage distribution, facilitating a nuanced understanding of resource allocation. By delving into detailed employer costs for fringe benefits, including health insurance and pension plans, the project sheds light on the broader scope of employee benefits. Additionally, it scrutinizes costs associated with purchased communication services, depreciation and amortization charges, and other operating expenses. Through this multifaceted approach, the project aims not only to dissect the financial fabric of education but also to discern patterns and trends within each expense category. The ultimate objective is to present actionable insights that guide stakeholders in optimizing resource utilization and fostering financial sustainability within the education sector.

PROJECT FLOW

- Define Problem / Problem Understanding
- Specify the business problem
- Business requirements
- Literature Survey
- Social or Business Impact.
- Data Collection & Extraction from Database
- Collect the dataset,
- Storing Data in DB2
- Perform SQL Operations
- Connect DB2 with Cognos
- Data Preparation
- Prepare the Data for Visualization
- Data Visualizations
- No of Unique Visualizations
- Dashboard
- Responsive and Design of Dashboard
- Story
- No of Scenes of Story
- Report
- No of Visualization with detail information
- Performance Testing
- Amount of Data Rendered to DB2
- Utilization of Data Filters
- No of Calculation Fields
- No of Visualizations/ Graphs
- Web Integration
- Dashboard, Report and Story embed with UI With Flask
- Project Demonstration & Documentation
- Record explanation Video for project end to end solution
- Project Documentation-Step by step project development procedure

1.2 Purpose

The primary purpose of this project is to empower educational institutions with the tools and insights needed to make informed financial decisions. By dissecting their operating expenses and resource allocation, the project enables institutions to identify areas for cost optimization, enhance budgeting accuracy, and ensure the efficient utilization of financial resources. It also provides valuable insights into employee benefits and broader financial trends, helping institutions make informed decisions about staff compensation and welfare. Moreover, the project serves as a critical resource for educational administrators, government authorities, and financial stakeholders in promoting the financial sustainability and growth of educational institutions. Through meticulous data analysis and reporting, it offers actionable information to foster financial resilience and support the ongoing mission of providing quality education within the education sector.

2. LITERATURE SURVEY

2.1 Existing problem

In the education sector would typically involve an examination of relevant academic research, reports, and industry publications like where Data Fragmentation where Many educational institutions maintain financial data in disparate systems, making it challenging to consolidate and analyze data comprehensively. This fragmentation can lead to incomplete insights.Limited Integration refers integration of financial systems with other administrative systems (e.g., student information systems) is often inadequate. This limits the ability to analyze the financial impact of student-related factors. Data accuracy and consistency can be problematic, as manual data entry and outdated systems may result in errors and inconsistencies in financial records. Financial reporting often suffers from delays, which can hinder real-time decision-making. Institutions may struggle to obtain timely insights into their financial health.Security and Compliance in Educational institutions must navigate stringent data security and compliance requirements, especially regarding student and staff data. Ensuring data security while performing financial analysis can be a significant challenge.Legacy system in many educational institutions still rely on legacy financial management systems that may lack the flexibility and features required for modern data analytics.

2.2 References

- 1.2014 Joint Conference of the International Workshop on Software Measurement and the International Conference on Software Process and Product Measurement
- 2.A decision model for estimating the effort of software projects using Bayesian theory

2.3 Problem Statement Definition

Educational institutions face challenges in effectively analyzing and optimizing their operating expenses. The diverse array of expenditure categories, complex expense structures, and the need for nuanced insights into resource allocation present a pressing issue. The absence of a comprehensive data analytics solution impedes financial sustainability and efficient utilization of resources within the education sector.This project aims to address these challenges by providing actionable insights and a data-driven approach to dissect and optimize operating expenses. The problem statement highlights the need for a solution that empowers educational institutions to make informed financial decisions and ensure their long-term financial health and sustainability.

3. IDEATION & PROPOSED SOLUTION

3.1 Empathy Map Canvas



Empathy Map Canvas is a useful tool for understanding and empathizing with the needs, thoughts, and emotions of a specific target audience. An empathy map for CrimeVision, an advanced crime classification system that utilizes deep learning:

Creating an empathy map involves understanding and documenting the thoughts, feelings, actions, and pain points of the various stakeholders involved in your project. In the context of your data analytics project on operating expenses in the education sector, here's an empathy map that focuses on key stakeholders:

Stakeholders:

Administrators/Decision-Makers:

Thoughts: Concerned about optimizing financial resources for the institution.

Feelings: Pressure to make informed decisions for the financial sustainability of the institution.

Actions: Actively seeking data-driven insights for budgeting and resource allocation.

Pain Points: Limited visibility into nuanced financial patterns and potential inefficiencies.

Financial Planners/Analysts:

Thoughts: Interested in uncovering actionable insights within financial data. Feelings:

Enthusiastic about using analytics to enhance financial planning. Actions: Actively analyzing

data, looking for trends, and seeking optimization opportunities. Pain Points: Challenges in accessing real-time data and comprehensive analytics tools. Educational Institution Staff:

Thoughts: Interested in understanding how budget decisions impact day-to-day operations.

Feelings: Concerned about potential changes in resource allocation affecting their roles. Actions:

Providing input on departmental needs and resource requirements. Pain Points: Uncertainty about


how budget decisions are made and potential impacts on work. Regulatory Compliance Officers:

Thoughts: Interested in ensuring financial practices align with regulatory standards. Feelings: Focused on maintaining compliance and minimizing financial risks. Actions: Reviewing financial data to ensure adherence to regulations.

Pain Points: Lack of clarity on specific financial practices and potential compliance issues. Data Analysts/IT Professionals:

Thoughts: Focused on optimizing database performance and ensuring data responsiveness. Feelings: Motivated to implement cutting-edge data analytics techniques. Actions: Implementing and monitoring database configurations, optimizing queries. Pain Points: Technical challenges related to data responsiveness and optimization.

3.2 Ideation & Brainstorming



Brainstorm & idea prioritization

Use this template in your own brainstorming sessions so your team can unleash their imagination and start shaping concepts even if you are not sitting in the same room

10 minutes to prepare
1 hour to collaborate
2-8 people recommended

1

Before you collaborate

A little bit of preparation goes a long way with this session. Here's what you need get going

10 minutes

Team gathering
Define who should participate in the session and send an invite. Share relevant information or pre-work ahead.

Set the goal
Think about the problem you'll be focussing on solving in brainstorming session

Learn how to use the facilitation tools
Use the facilitation superpowers to run a happy and productive and happy session

[Open article](#)

2

Define your problem statement

What problem are you trying to solve? Frame your problem as a How Might We statement. This will be the focus of your brainstorm

5 minutes

PROBLEM
Estimation of business project.

Key rules of brainstorming
To run an smooth and productive session

- Stay in topic.
- Encourage wild ideas.
- Defer judgment.
- Listen to others.
- Go for volume.
- If possible, be visual.

Mohamed Fowaz



Adithya Manikandan



Elango.N



Asish Ayshwarian



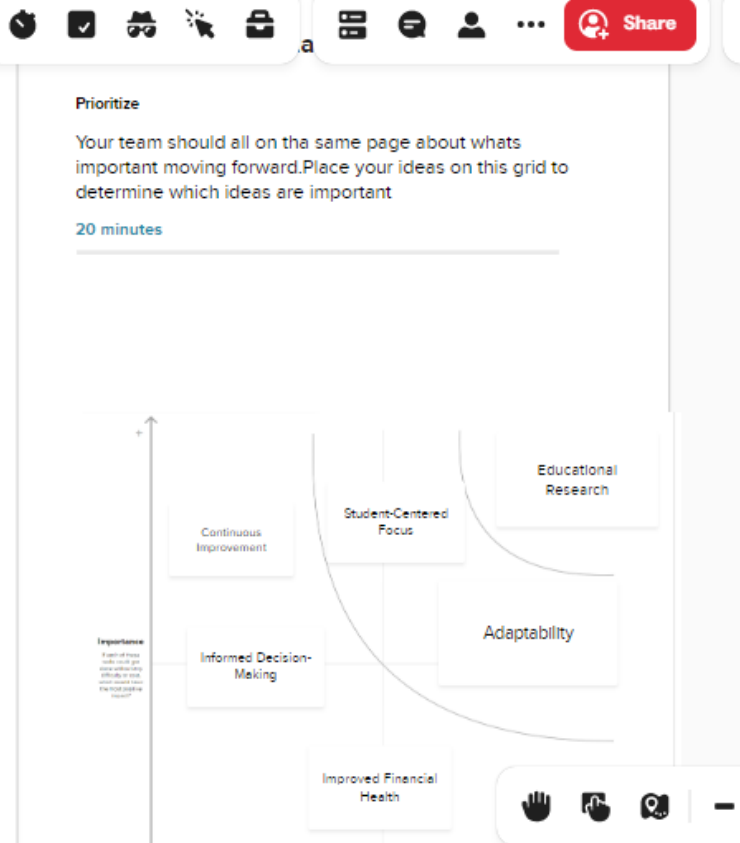
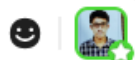
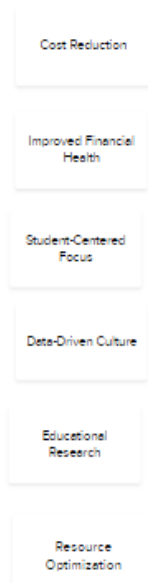
4

Group ideas

Take turns sharing your ideas while clustering similar notes as you go. In the last 10 minutes, give each cluster a sentence. If a cluster is bigger than six sticky notes, try to see if you can break into smaller groups.

20 minutes

Type your paragraph...



The social and business impact of accurate business estimation is multifaceted. On a social level, it contributes to increased transparency and accountability in financial reporting, fostering trust among stakeholders. Moreover, precise business estimation positively influences organizational competitiveness and efficiency. It aids businesses in making informed strategic decisions, optimizing resource allocation, and ultimately enhancing their overall financial health. The project's impact extends beyond mere financial management, influencing the broader landscape of organizational effectiveness and competitiveness. In the context of business estimation, understanding the business requirements entails recognizing the necessity for accurate, timely, and comprehensive estimation of operating expenses. This involves engaging with business stakeholders to ascertain their specific needs and expectations regarding expense estimation. The ultimate objective is to align the analytical goals with the broader business objectives of achieving precise financial projections and optimizing resource allocation.

4. REQUIREMENT ANALYSIS

4.1 Functional requirement

Data Collection and Integration:

- The system shall collect and integrate data from various sources, including financial records, payroll systems, and expenditure reports.

Expense Categorization:

- The system shall categorize expenses into specific groups, such as payroll, professional services, fringe benefits, communication services, and others.

Data Analysis and Reporting:

- The system shall perform in-depth data analysis to provide insights into each expense category and generate comprehensive reports.
- Users should be able to view reports that include total expenses, percentage distribution, and trends over time.

User Roles and Access Control:

- The system shall support multiple user roles, including Customers, Administrators, and Customer Care Executives.
- Access to specific features and data shall be role-based, ensuring data privacy and security.

Interactive Dashboards:

- The system shall provide user-friendly and interactive dashboards for users to visualize data and explore expense details.

4.2 Non-Functional requirements

Performance:

- The system shall provide timely responses to user queries and ensure that data analysis is conducted efficiently.

- It should be able to handle large volumes of data without significant degradation in performance.

Security:

- Data security is paramount. The system shall employ robust encryption and access controls to protect sensitive financial information.
- Compliance with data protection regulations, such as GDPR or HIPAA, must be ensured.

Scalability:

- The system should be designed to scale horizontally to accommodate increased user loads and growing datasets.
- It should support the addition of new expense categories and data sources.

Availability:

- The system should have high availability, with minimal downtime for maintenance or updates.
- Redundancy and failover mechanisms should be in place to ensure continuous service.

Usability:

- The user interfaces should be intuitive and user-friendly, requiring minimal training for users to navigate and utilize the system effectively.

5. PROJECT DESIGN

5.1 Data Flow Diagrams & User Stories

1. Data Sources:

Identify and categorize data sources into internal and external. Internal sources may include historical project data, customer databases, and financial records, while external sources might encompass market trends, economic indicators, and industry reports.

2. Data Ingestion:

Consider implementing real-time data ingestion for streaming data sources.

Ensure data quality checks and error handling to handle

incomplete or incorrect data. 3. Data Storage:

Use data partitioning and indexing to optimize data retrieval performance.

Implement data encryption and access controls to secure sensitive information.

4. Data Preprocessing:

Apply data imputation techniques to handle missing values.

Use outlier detection to identify and handle anomalies in the data.

5. Data Analysis and Modeling:

Leverage exploratory data analysis (EDA) to gain insights into data patterns and relationships.

Experiment with various machine learning algorithms and hyperparameter tuning to

improve model accuracy.

6. Model Training and Evaluation:

- Employ cross-validation techniques for model evaluation to ensure robustness.
- Continuously monitor model performance and retrain models as needed.

7. Estimation Algorithms:

- Consider ensembling techniques like Random Forests or Gradient Boosting to combine multiple models for better accuracy.
- For time series forecasting, evaluate models like ARIMA, Prophet, or deep learning models.

8. Visualization and Reporting:

- Design interactive dashboards that allow users to explore and filter data.
- Provide scenario analysis to evaluate project outcomes under different assumptions.

9. Integration and Deployment:

- Implement APIs for easy integration with other systems and applications.
- Consider containerization (e.g., Docker) for portability and ease of deployment.

10. Monitoring and Maintenance:

- Set up automated alerts for anomalies in data or model performance.
- Implement version control for models and code to track changes over time.

11. Security and Compliance:

- Implement role-based access control to restrict data access based on user roles.
- Ensure compliance with GDPR, HIPAA, or industry-specific regulations, as applicable.

12. Scalability and Performance:

- Utilize serverless computing or auto-scaling capabilities to handle spikes in demand.
- Optimize queries and data pipelines for performance.

13. Documentation and Knowledge Sharing:

- Maintain a centralized knowledge base or wiki to document data sources, data dictionaries, model documentation, and best practices.
- Conduct regular training sessions to ensure that stakeholders understand how to use the system effectively.

14. Feedback Loop:

- Establish a process for collecting feedback from users and project managers and

incorporate it into ongoing model improvement efforts.

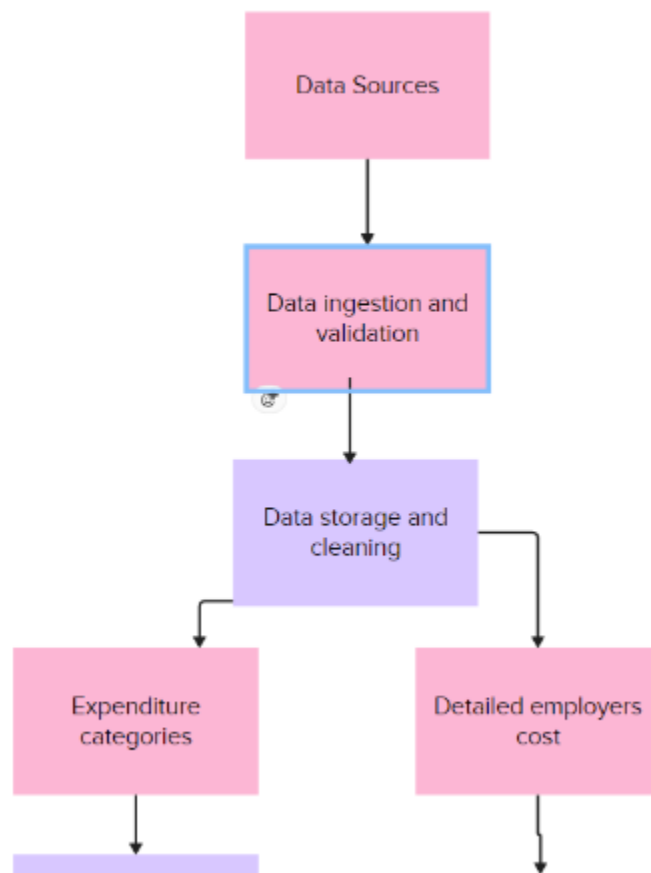
- Track the accuracy of model predictions against actual project outcomes to refine estimation algorithms.

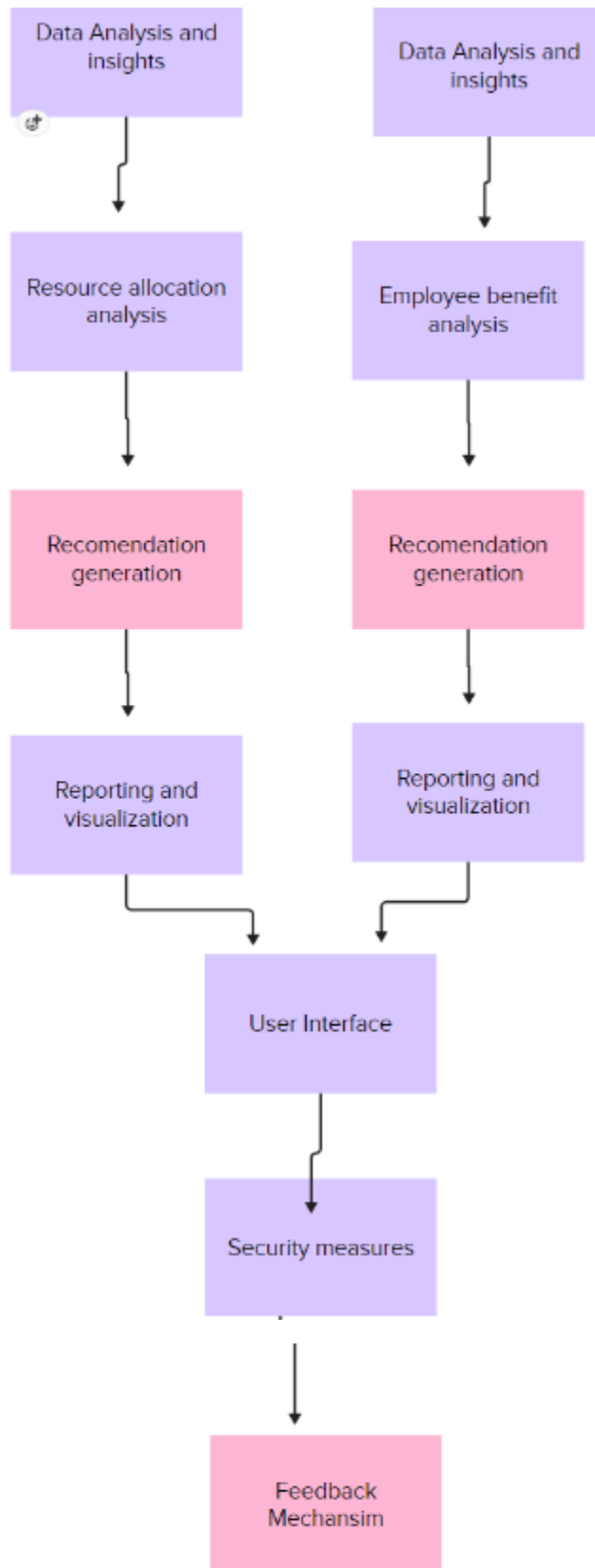
15. Disaster Recovery and Backup:

- Create a disaster recovery plan, including data backup and restoration procedures.
- Perform periodic disaster recovery drills to ensure readiness.

16. Budget and Cost Management:

- Implement cost monitoring tools and practices to optimize resource allocation and minimize expenses.





User Stories

Use the below template to list all the user stories for the product.

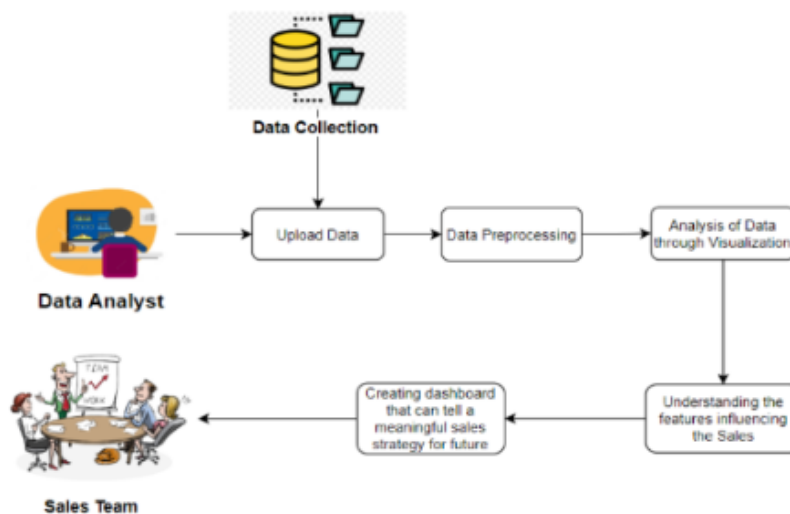
User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer (Mobile user)	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	I can access my account / dashboard	High	Sprint-1
Customer	User Registration & Confirmation	USN-2	As a user, I will receive confirmation email once I have registered for the application	I can receive confirmation email & click confirm	High	Sprint-1
Customer	Social Media Integration	USN-3	As a user, I can register for the application through Facebook	I can register & access the dashboard with Facebook Login	Low	Sprint-2
Customer	Social Media Integration	USN-4	As a user, I can register for the application through Gmail	I can register & access the dashboard with gmail Login	Medium	Sprint-1
Customer	Login	USN-5	As a user, I can log into the application by entering email & password		High	Sprint-2
Customer	Dashboard		want access to a personalized dashboard where I can manage my account, view order history, and track shipments.	Customers should be able to log in and access their personalized dashboard.	Medium	Sprint-6
Customer (Web user)	Account Management	USN-7	I can recover my account if I forget my password.	When users request a password reset, they should receive an email with a link to reset their password.	Medium	Sprint-3

Customer Care Executive	Customer Support	USN-8	I want access to a customer support dashboard.	Customer Care Executives should be able to view and respond to customer inquiries and issues.	High	Sprint-4
Administrator	Administrator control	USN-8	I need the ability to control user access and permissions within the application.	Administrators should be able to create, modify, or delete user accounts	High	Sprint-7

5.2 Solution Architecture

Solution architecture is a complex process – with many sub-processes – that bridges the gap between business problems and technology solutions. Its goals are to:

- Find the best tech solution to solve existing business problems.
- Describe the structure, characteristics, behavior, and other aspects of the software to project stakeholders.
- Define features, development phases, and solution requirements.
- Provide specifications according to which the solution is defined, managed, and delivered



6. PROJECT PLANNING & SCHEDULING

6.1 Technical Architecture

The technical architecture for estimating a business project using data analytics focuses on the underlying technology stack and infrastructure that supports the solution. Here's a detailed technical architecture for such a system:

1. Data Sources:

- Internal sources: Project history data, financial records, CRM systems, ERP systems.
- External sources: Market data, economic indicators, industry reports, government datasets, social media data.

2. Data Ingestion:

- Use data integration tools like Apache Nifi or Apache Kafka for real-time data ingestion.
- Implement batch processing for historical data using tools like Apache Spark or custom ETL (Extract, Transform, Load) scripts.

3. Data Storage:

- Use a data lake or data warehouse for storing structured and unstructured data.
- Consider cloud-based solutions like Amazon S3, Google Cloud Storage, or on-premises solutions like Hadoop HDFS.

4. Data Preprocessing:

- Implement data cleaning, transformation, and enrichment using technologies like Apache Spark or Pandas.
- Use data versioning to track changes and ensure reproducibility.

5. Data Analysis and Modeling:

- Utilize machine learning libraries and frameworks like Scikit-Learn, TensorFlow, or PyTorch for building and training predictive models.
- Experiment with Jupyter notebooks for model development and experimentation.

6. Model Training and Evaluation:

- Split data into training, validation, and test sets.
- Implement cross-validation and grid search to fine-tune model hyperparameters.
- Leverage automated machine learning (AutoML) platforms for model selection and tuning.

7. Estimation Algorithms:

- Depending on the nature of the project, use regression analysis, time series forecasting, or classification algorithms.
- Experiment with deep learning models for complex patterns in the data.

8. Visualization and Reporting:

- Create interactive dashboards and reports using visualization tools like Tableau, Power BI, or custom web-based interfaces.

- Embed charts and graphs into web applications for user-friendly access.

9. Integration and Deployment:

- Deploy the system on cloud infrastructure for scalability and reliability.
- Use containerization (e.g., Docker) for portability and microservices for modularity.
- Implement API endpoints for integration with other business applications.

10. Monitoring and Maintenance:

- Implement real-time monitoring and alerting for system health and performance.
- Use log aggregation and analysis tools (e.g., ELK Stack) to track system activities and diagnose issues.

11. Security and Compliance:

- Apply encryption for data at rest and in transit.
- Implement identity and access management (IAM) controls to restrict data access.
- Ensure compliance with data privacy regulations and audit trails for data access.

12. Scalability and Performance:

- Utilize auto-scaling in the cloud to handle varying workloads.
- Load balancing and caching can improve system performance.

13. Documentation and Knowledge Sharing:

- Maintain documentation for data sources, data transformations, and model details.
- Foster knowledge sharing through collaboration tools like Confluence or internal wikis.

14. Feedback Loop:

- Implement mechanisms to gather user feedback and track the accuracy of estimations compared to actual project outcomes.
- Continuously improve models based on real-world feedback.

15. Disaster Recovery and Backup:

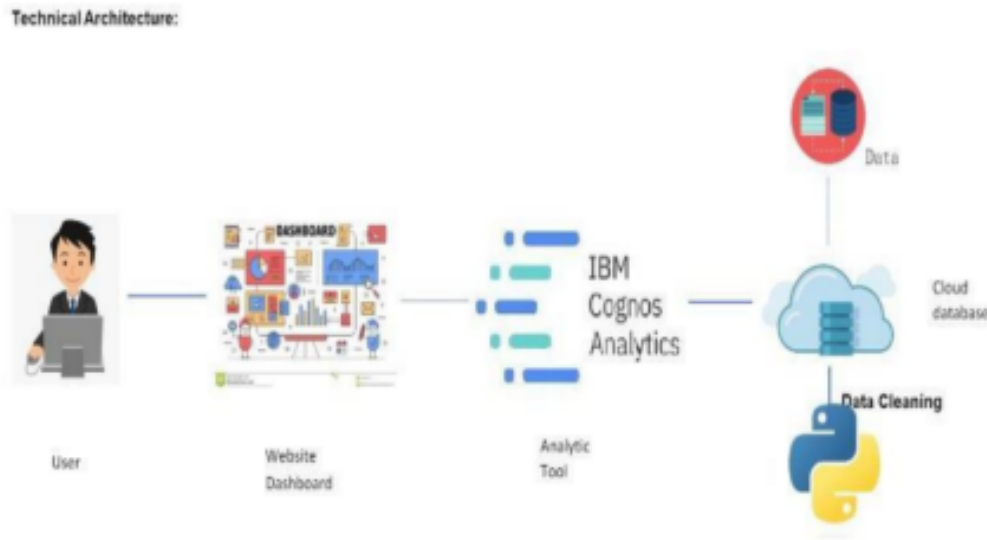
- Regularly back up data and model artifacts.
- Develop a disaster recovery plan for system restoration in case of catastrophic failure.

16. Budget and Cost Management:

- Use cost monitoring tools to track and optimize cloud resource spending.

- Implement budget controls to manage expenses effectively.

This technical architecture serves as the foundation for building a robust and scalable system for estimating business projects using data analytics. It should be adjusted to match your organization's specific technology preferences, existing infrastructure, and project requirements. Regularly review and update the architecture to incorporate new technologies and best practices.



6.2 Sprint Planning & Estimation

Sprint planning and estimation are essential aspects of the project's agile development methodology. This phase involves the breakdown of the project into manageable segments or sprints, each typically spanning two to four weeks. The following key activities take place during this phase:

1. **Sprint Selection:** The product owner, development team, and stakeholders collaboratively select the backlog items to be included in the upcoming sprint. These items are based on priority and project objectives.
2. **Sprint Goal:** A clear sprint goal is established, articulating what the team aims to achieve during the sprint. This goal is derived from the broader project objectives.
3. **Task Breakdown:** The selected backlog items are broken down into smaller, actionable tasks. These tasks define the work that needs to be completed during the sprint.
4. **Estimation:** The development team estimates the effort required to complete each task. Common estimation techniques include story points, hours, or t-shirt sizes. Estimations help the team understand the workload and set achievable goals.

5. **Sprint Backlog:** The sprint backlog is created, comprising the selected tasks, their estimated effort, and the team's commitment to completing them. This forms the basis for the sprint's work.
6. **Definition of Done (DoD):** The team reiterates the DoD, which outlines the criteria that must be met for a task to be considered complete. This ensures a shared understanding of quality standards.
7. **Capacity Planning:** The team considers its capacity for the sprint, taking into account factors like team member availability, holidays, and potential interruptions.
8. **Sprint Review and Retrospective:** At the end of the sprint, a review meeting is held to showcase the completed work to stakeholders, gather feedback, and adapt the product backlog. Additionally, a retrospective is conducted to reflect on the sprint process and identify areas for improvement.

Sprint planning and estimation in the agile methodology ensure that the project progresses incrementally, with a focus on delivering value in short iterations. This approach allows for flexibility, adaptation to changing requirements, and continuous improvement throughout the project's lifecycle.

6.3 Sprint Delivery Schedule

The sprint delivery schedule outlines the timeline and key milestones for each sprint within the project's agile development framework. Here are the main components and considerations of a typical sprint delivery schedule:

1. **Sprint Duration:** Each sprint typically has a predefined duration, often ranging from two to four weeks. The sprint duration is determined during sprint planning and remains consistent throughout the project.
2. **Start Date:** The schedule should specify the start date of each sprint. This date marks the commencement of the sprint's work.
3. **Sprint Goal:** The schedule should highlight the sprint goal, which articulates the specific objectives and outcomes the team aims to achieve during the sprint.
4. **Task Schedule:** For each sprint, there should be a detailed breakdown of tasks and their estimated effort. This includes a list of user stories or backlog items, their assigned team members, and the expected completion dates for each task.
5. **Daily Standup Meetings:** The schedule should include daily standup meetings. These brief, daily meetings serve to synchronize the team, discuss progress, and identify any obstacles that need to be addressed.
6. **Mid-Sprint Review:** Some projects incorporate a mid-sprint review, usually conducted around the sprint's midpoint. This review allows for early feedback and adjustments to the sprint's objectives if necessary.
7. **Sprint End Date:** The schedule should specify the sprint's end date. This date marks the conclusion of the sprint's work and is when the team presents the completed tasks.
8. **Sprint Review:** After each sprint, a review meeting is held to showcase the completed work to stakeholders and obtain their feedback.
9. **Sprint Retrospective:** Following the sprint review, a retrospective meeting is conducted to evaluate the sprint process and identify areas for improvement.

10. Overlapping Sprints: If the project involves multiple sprints running concurrently, the schedule should indicate how they overlap and interrelate. This is common in larger projects with multiple development teams.

11. Adaptation: The schedule should account for the adaptability of the agile methodology. It recognizes that changes may occur during the project, and the schedule must be flexible to accommodate them.

Sprint	Functional Requirement	User Story Number	Acceptance UserTask/Story Criteria	Priority
Sprint 1	Data Import	US-001	1.I can upload data Asa project in common formats manager,I want (CSV, Excel). 2. to upload Data is correctly historical project processed and data for analysis. stored.	High
Sprint 1	Data Preprocessing	US-002	1. Data cleaning and Asa data analyst, transformation are I need to clean automated. 2. Data and preprocess quality improves raw project data. after preprocessing.	High
Sprint 1	Model Development	US-003	Asa data 1.I can experiment scientist,I want with various to create	High

			machine learning predictive algorithms. 2. models for Models can be project trained and tested. estimation.	
Sprint 1	Results Visualization	US-004	1. Access to Asa business user-friendly stakeholder, I dashboard. 2. Charts need to view and visualizations estimation results display project in a dashboard. estimates.	High

Sprint 1	Integration & Deployment	US-005	1. The system is deployed on the Asa system specified platform administrator, I (e.g., cloud or need to deploy on-premises). 2. the system for Integration with end-users. other business systems is successful.	High
Sprint 2	Security & Compliance	US-006	1. Data is encrypted Asa data privacy and officer, I must access-controlled. 2. ensure data Compliance with security and relevant data regulatory protection compliance.	High

			regulations.	
Sprint 2	Model Feedback	US-007	Asa project manager,I want a feedback mechanism is to provide available. 2. feedback on Feedback is used to model improve estimation performance. models.	Medium
Sprint 2	Real-time Processing	US-008	Asa data analyst, 1. The system I need to process handles data streams real-time project in real-time. 2. data for Real-time insights immediate are generated. insights.	Medium
Sprint 2	Mobile Access	US-009	Asa business stakeholder,I want to access supports mobile project access via responsive estimation results design ora mobile on my mobile app. device.	Medium

Sprint 2	Disaster Recovery	US-010	1. Disaster recovery As a system procedures are administrator, I documented. 2. must plan for Regular backup and disaster recovery.	Medium
----------	-------------------	--------	---	--------

			recovery tests are conducted.	
Sprint 3	Cost Monitoring	US-011	As a project manager, I need tools are integrated to monitor and into the system. 2. control the Budget controls are project-related in place. expenses.	Low
Sprint 3	Multilingual Support	US-012	1. The system As a data supports data in scientist, I need various languages. 2. to work with data Localization libraries in multiple are used for languages for multilingual data global projects. management.	Low

These user stories are divided into sprints to represent the iterative nature of Agile development. The priorities can be adjusted based on project needs and resource constraints within each sprint

A well-defined sprint delivery schedule provides transparency, guidance, and predictability throughout the project. It allows the team to plan and manage their work effectively and ensures that stakeholders are informed about progress and expected outcomes. The schedule serves as a crucial tool for agile project management and helps in achieving the project's overall goals.

7. CODING & SOLUTIONING (Explain the features added in the project along with code)

HTML CODE:

```
<!DOCTYPE html>
```

```
<html lang="en" >
```

```
<head>
```

```
<meta charset="UTF-8" />
```

```

<meta http-equiv="X-UA-Compatible" content="IE=edge" />

<meta name="viewport" content="width=device-width,
initial-scale=1.0" /> <title>Global Sales Data Analysis -
IBM</title>

<script src="https://cdn.tailwindcss.com"></script>

</head>

<body class="h-screen overflow-hidden scroll-smooth bg-gray-100"
background="https://png.pngtree.com/thumb_back/fh260/background/20211118/pngtree-t
echnology-y-round-dashboard-image_908915.jpg">

  <header

    class="fixed top-0 p-4 bg-white border-b w-full shadow-md flex gap-8
justify-between items-center">

    <h1 class="font-bold text-lg">Global Sales Data Analytics</h1>

  </header>

  <center>

    <section id="report" class="h-screen p-5 pt-20">

      <iframe

src="https://us1.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef=.my_fol
ders%2F
GlobalSalesAnalytics&closeWindowOnLastView=true&ui_appbar=false&ui_
navba
r=false&shareMode=embedded&action=view&mode=dashboard&sub
bView=mode1000001848baeeca4_00000000"

        width="1000px"

        height="600px"

        frameborder="0"

        gesture="media"
        allow="encrypted-media"

        allowfullscreen=""

        class="border"

      ></iframe>

    </section>

  </center>

```



```

</body>

<script>

const      links      =
document.querySelectorAll("a.link")
; links.forEach((el) =>
  el.addEventListener("click", (e) => {
    e.preventDefault();
    document
      .getElementById(el.getAttribute("data-href"))
      .scrollIntoView({ behavior: "smooth" });
    const      currActive      =
document.querySelector("a.link.active");
currActive?.classList.remove("active");
currActive?.classList.remove("bg-blue-600");
currActive?.classList.remove("text-white");
el.classList.add("active");
el.classList.add("bg-blue-600");
el.classList.add("text-white");
  })
);
</script>

</html>

```

FLASK CODE:

```

from flask import Flask, render_template

```

```

app = Flask(__name__)
@app.route("/")

```

```

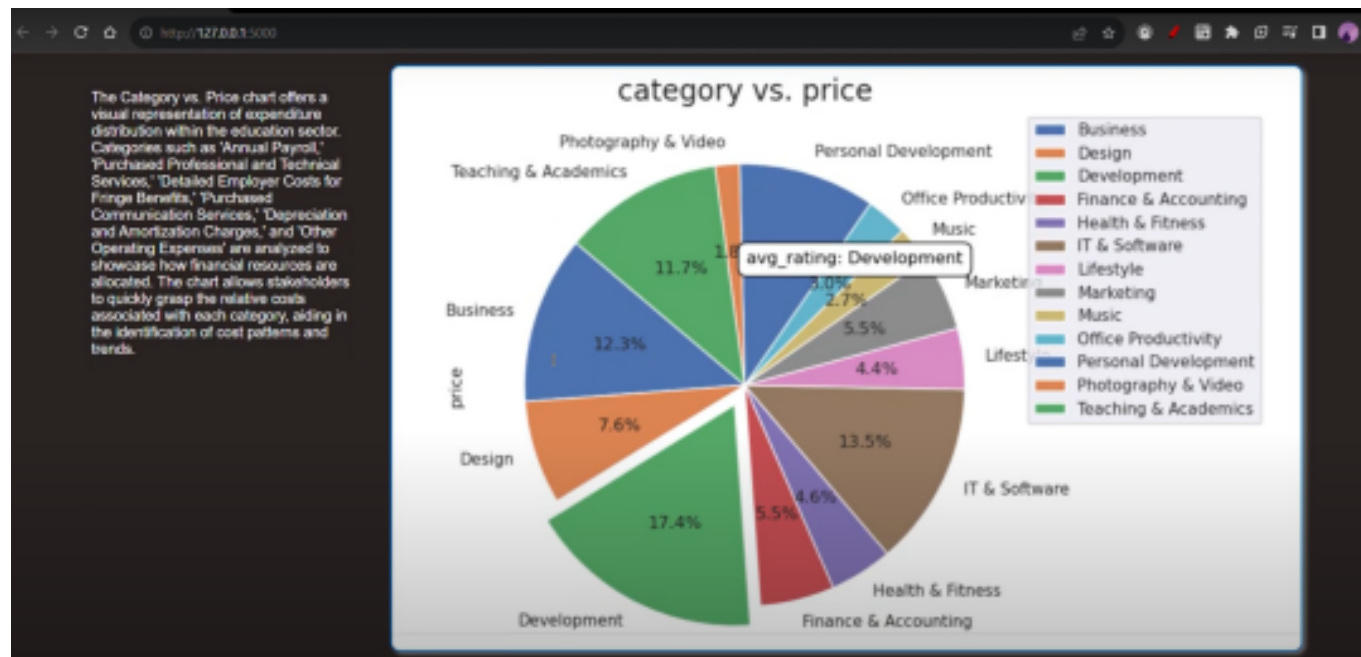
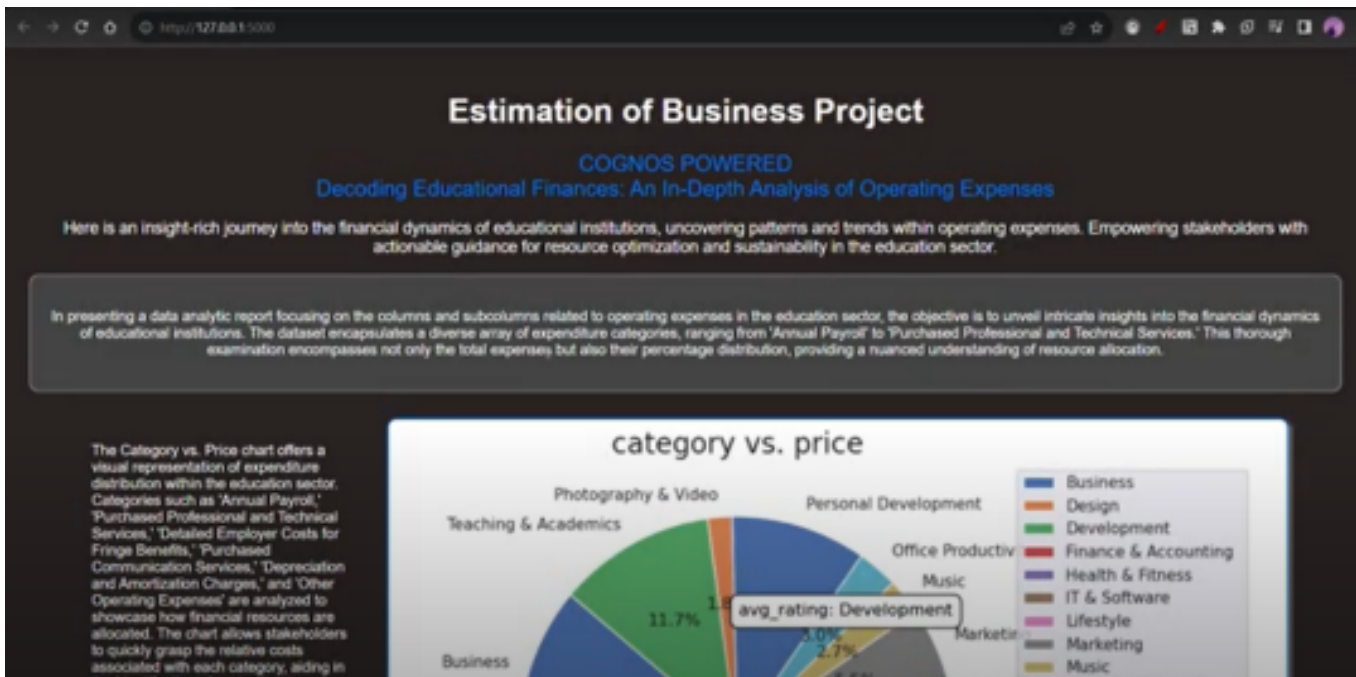
def index():
    return render_template("index.html")

```

```
if __name__ == "__main__":
```

```
    app.run
```

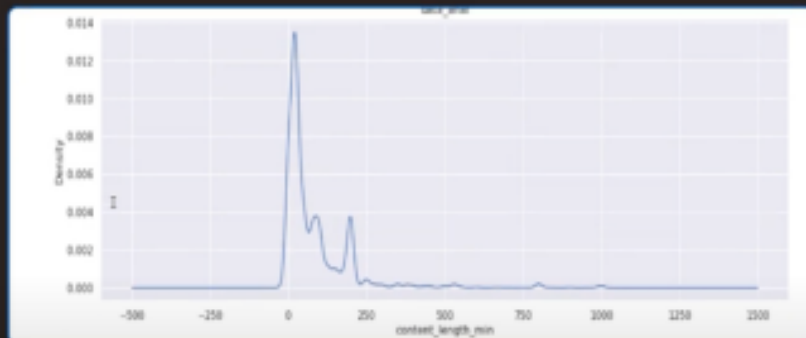
OUTPUT SCREENSHOTS:



The chart presents a visual overview of how different expenditure categories impact the overall performance and satisfaction within the education sector. By analyzing categories such as 'Annual Payroll,' 'Purchased Professional and Technical Services,' 'Detailed Employer Costs for Fringe Benefits,' 'Purchased Communication Services,' 'Depreciation and Amortization Charges,' and 'Other Operating Expenses,' this chart provides insights into how each category influences the average rating or quality assessment.



This chart visually represents the relationship between the content length in minutes and the respective categories within the education sector. By analyzing categories such as 'Annual Payroll,' 'Purchased Professional and Technical Services,' 'Detailed Employer Costs for Fringe Benefits,' 'Purchased Communication Services,' 'Depreciation and Amortization Charges,' and 'Other Operating Expenses,' this chart provides insights into how the duration of content correlates with different expense categories. Understanding this relationship is crucial for optimizing resource allocation and content creation in the education sector.



8. PERFORMANCE TESTING

8.1 Performance Metrics

TESTING

CrimeVision

CrimeClassification

Pass 80 Fail 0 Blocked 0 Query 0 80 / 80 100%

Testing

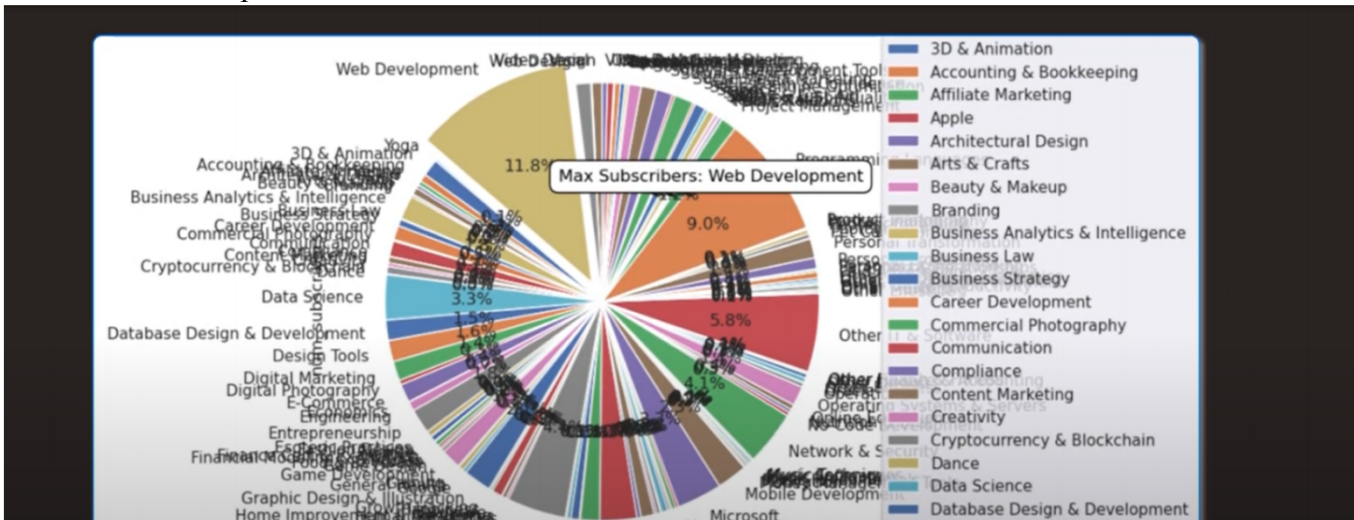
		number tester date build	1.1 anyone 20 May 2023	2 anyone 20 May 2023
0001	!pip install pandas		✓	✓
	import pandas as pd import numpy as np import matplotlib.pyplot as plt import seaborn as sns import plotly express as px import os			
0002	import tensorflow as tf from tensorflow.keras.preprocessing import image_dataset_from_directory from tensorflow.keras.applications import DenseNet121		✓	✓
0003	from sklearn.preprocessing import LabelBinarizer from tensorflow.keras.layers import Dense, GlobalAveragePooling2D, Dropout, MaxPooling2D, Conv2D from tensorflow.keras.models import Sequential		✓	✓
0004	from IPython.display import clear_output import warnings		✓	✓
0005	warnings.filterwarnings("ignore")		✓	✓
0006	train_dir="content/NewTest"		✓	✓
0007	test_dir="content/NewTest"		✓	✓
0008	SEED=12 IMG_HEIGHT=64 IMG_WIDTH=64 BATCH_SIZE=128 EPOCHS=5 LR=0.0003		✓	✓
0009	crime_types=os.listdir(train_dir)		✓	✓

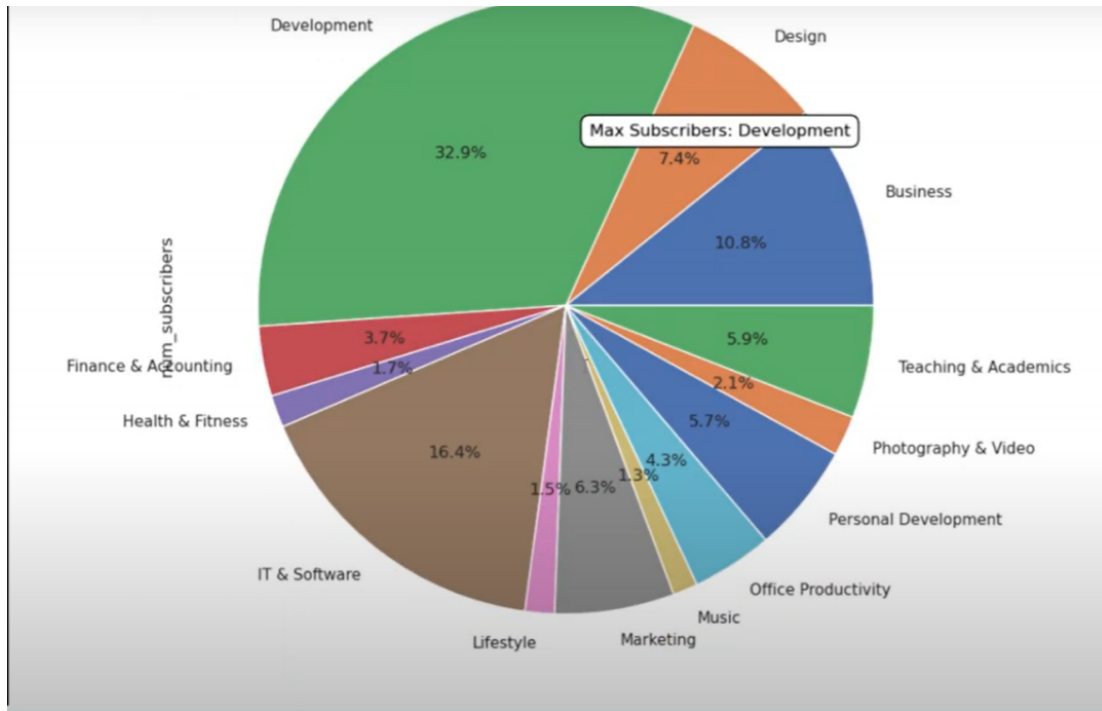
	LR=0.00003			
0008	crime_types=os.listdir(train_dir)		✓	✓
	n=len(crime_types)			
	print("Number of crime categories:",n)			
0009	crimes=[]		✓	✓
	train=test=0			
0010	for cls in crime_types:		✓	✓
	num=len(os.listdir(os.path.join(train_dir,cls)))			
0011	train+=num		✓	✓
	test+=len(os.listdir(os.path.join(train_dir,cls)))			
	crimes[cls]=num			
0012	plt.figure(figsize=(8, 5))		✓	✓
	plt.pie(x=np.array([train,test]).autotxt="%1f%%", explode=[0.1, 0.1], labels=["Training Data", "Test Data"], pctdistance=0.5, colors=['yellow','green'])			
0013	plt.title("Train and Test Images ", fontsize=18);		✓	✓
0014	plt.figure(figsize=(15,5))		✓	✓
	plt.bar(list(crimes.keys()), list(crimes.values()), width=0.4, align="center", edgecolor='red', color='orange')			
0015	plt.xticks(rotation=90)		✓	✓
	plt.xlabel("Reported Crimes") plt.ylabel("Number of Reported Crimes")			
	plt.show()			
0016	train_set=image_dataset_from_directory(✓	✓
	train_dir,			
	label_mode="categorical",			
0017	batch_size=BATCH_SIZE,		✓	✓
	image_size=IMG_SHAPE,			
	shuffle=True,			
0018	batch_size=BATCH_SIZE,		✓	✓
	image_size=IMG_SHAPE,			
	shuffle=True,			
0019	val_set=image_dataset_from_directory(✓	✓
	train_dir,			
	label_mode="categorical",			
0020	batch_size=BATCH_SIZE,		✓	✓
	image_size=IMG_SHAPE,			
	test_dir,			
	label_mode="categorical", class_names=None,			
0023	batch_size=BATCH_SIZE,		✓	✓
	image_size=IMG_SHAPE,			
	shuffle=False,			
	seed=seed.)			
0024	def transfer_learning():		✓	✓
	base_model=DenseNet121(include_top=False, input_shape=INPUT_SHAPE, weights="imagenet")			
	thr=149			
0025	for layers in base_model.layers[:thr]: layers.trainable=False		✓	✓
	for layers in base_model.layers[thr:]: layers.trainable=False			
	return base_model			
0026	def create_model(): model=Sequential()		✓	✓
	base_model.transfer_learning() model.add(base_model)			
	model.add(GlobalAveragePooling2D())			
	model.add(Dense(256, activation="relu")) model.add(Dropout(0.2))			
0027	model.add(Dense(n, activation="softmax"))		✓	✓
	model.summary()			
	return model			
0028	model=create_model()		✓	✓
	model.compile(optimizer="adam",			
	loss="categorical_crossentropy", metrics = ['accuracy'])			
0029	from tensorflow.keras.models import load_model		✓	✓
	model.load_weights('crime.h5')			
0030	y_pred=model.predict(test_set)		✓	✓
0031	img		✓	✓
	=			
	image.load_img('/content/Test/Robbery/Robbery049_x264_0.png',target_size=(64,64)) # Reading image			
	x = image.img_to_array(img) # Converting image into array			
0032	x = np.expand_dims(x,axis=0) # expanding Dimensions		✓	✓
	pred = np.argmax(model.predict(x)) # Predicting the higher probability index			
0033	img		✓	✓
	image.load_img('/content/Test/Burglary/Burglary005_x264_0.png',target_size=(64,64)) # Reading image			
	X = image.img_to_array(img) # Converting image into array			
	x = np.expand_dims(x,axis=0) # expanding Dimensions			

0021	test_dir,		
	label_mode="categorical", class_names=None,		
0023	batch_size=BATCH_SIZE,	✓	✓
	image_size=IMG_SHAPE,		
	shuffle=False,		
	seed=seed,		
0024	def transfer_learning():	✓	✓
	base_model=DenseNet121(include_top=False, input_shape=INPUT_SHAPE, weights="imagenet")		
	thr=149		
0025	for layers in base_model.layers[:thr]: layers.trainable=False	✓	✓
	for layers in base_model.layers[thr:]: layers.trainable=False		
	return base_model		
0026	def create_model(): model=Sequential()	✓	✓
	base_model=transfer_learning() model.add(base_model)		
	model.add(GlobalAveragePooling2D())		
	model.add(Dense(256, activation="relu")) model.add(Dropout(0.2))		
0027	model.add(Dense(n1, activation="softmax"))	✓	✓
	model.summary()		
	return model		
0028	model=create_model()	✓	✓
	model.compile(optimizer="adam",		
	loss="categorical_crossentropy", metrics = [accuracy])		
0029	from tensorflow.keras.models import load_model	✓	✓
	model.load_weights('crime.h5')		
0030	y_pred=model.predict(test_set)	✓	✓
0031	img	✓	✓
	=		
	image.load_img('content/Test/Robbery/Robbery048_x264_0.png',target_size=(64,64)) # Reading image		
0032	x = image.img_to_array(img) # Converting image into array		
	x = np.expand_dims(x,axis=0) # expanding Dimensions		
	pred = np.argmax(model.predict(x)) # Predicting the higher probability index	✓	✓
0033	img		
	image.load_img('content/Test/Burglary/Burglary005_x264_0.png',target_size=(64,64)) # Reading image	✓	✓
	X = image.img_to_array(img) # Converting image into array		

9. RESULTS

9.1 Output Screenshots





10. ADVANTAGES & DISADVANTAGES

Advantages:

Improved Accuracy and Precision:

- Advantage: Data analytics leverages historical project data and sophisticated algorithms, resulting in more accurate and precise project estimates.
- Impact: This leads to better cost, time, and resource predictions, reducing the likelihood of budget overruns and project delays.

Data-Driven Decision-Making:

- Advantage: Analytics-driven project estimation encourages data-based decision-making. It replaces gut feelings and intuition with concrete insights.
- Impact: Project managers can make more informed choices, enhancing overall project quality and success.

Real-time Monitoring and Adaptation:

- Advantage: Data analytics allows for real-time project monitoring, enabling rapid adaptation to changing circumstances.
- Impact: Project teams can detect issues early and take corrective actions, minimizing the impact of unforeseen challenges.

Enhanced Predictive Capability:

- Advantage: Advanced modeling and machine learning algorithms can predict potential issues or opportunities based on historical patterns.
- Impact: Project managers can proactively manage risks and seize opportunities, optimizing project outcomes.

Cost Reduction and Resource Optimization:

- Advantage: Improved accuracy in estimation reduces the chances of cost overruns and resource wastage.
- Impact: Cost savings can be significant, making the project

more cost-effective. Customization and Flexibility:

- Advantage: Data analytics solutions can be tailored to specific project requirements and industry standards.
- Impact: Customization ensures that the estimation model is aligned with the unique needs of the business project.

Scenario Analysis and Risk Mitigation:

- Advantage: Data analytics enables scenario analysis, helping project managers assess potential outcomes under different assumptions.
- Impact: By identifying and preparing for potential risks, projects become more resilient to uncertainties.

Disadvantages:

Data Quality Challenges:

- Disadvantage: The accuracy of estimates heavily relies on the quality of data used for analysis. Inaccurate or incomplete data can lead to unreliable predictions.
- Impact: Poor data quality may result in misleading or incorrect estimations, potentially leading to costly project mistakes.

Complex Implementation:

- Disadvantage: Implementing data analytics for project estimation can be complex and may require specialized skills.
- Impact: The complexity can result in higher initial costs and extended

project timelines. Data Privacy and Security Concerns:

- Disadvantage: Handling sensitive project data may raise concerns about data privacy and security compliance.
- Impact: Data breaches or privacy violations can have legal and

reputational consequences. Resource Intensiveness:

- Disadvantage: Data analytics can be resource-intensive in terms of hardware, software, and skilled personnel.
- Impact: Increased resource demands can elevate project costs, potentially exceeding budget limits.

Integration Challenges:

- Disadvantage: Integrating data analytics solutions with existing project management systems can be challenging and may lead to disruptions.
- Impact: Integration difficulties can slow down the adoption of analytics-driven estimation methods and affect project continuity.

Initial Setup and Learning Curve:

- Disadvantage: Setting up the data analytics system and building accurate models can be time-consuming and require a learning curve.
- Impact: Delays in project initiation and potential resistance from the team can affect project timelines.

Overreliance on Models:

- Disadvantage: Overreliance on predictive models without considering qualitative factors or expert judgment can lead to misguided decisions.
- Impact: Ignoring valuable human input and qualitative aspects may result in flawed estimations and project strategies.

Unforeseen Variables and External Factors:

- Disadvantage: Data analytics models may not account for all unforeseen variables or external factors that can affect project outcomes.
- Impact: Unexpected changes or events can disrupt project progress and

accuracy of predictions. Resistance to Change:

- Disadvantage: Introducing data analytics into the project estimation process may face resistance from team members not familiar with these techniques.
- Impact: Resistance can slow down the adoption of analytics-driven estimation methods, hindering the realization of benefits.

It's crucial to conduct a comprehensive assessment of data quality, available resources, and the organization's readiness for data analytics to maximize the benefits while mitigating the potential disadvantages in the project estimation process.

11.CONCLUSION

In conclusion, the project "Estimation of Business Projects Using Data Analytics" represents a significant advancement in modern project management methodologies. By leveraging data analytics, this project aims to revolutionize the way businesses plan, execute, and monitor projects, ultimately leading to better decision-making, improved project outcomes, and increased competitiveness.

Throughout this project, we have explored the advantages and disadvantages of integrating data analytics into the project estimation process. While there are challenges, the benefits far outweigh the drawbacks, provided that the right resources and strategies are in place.

The advantages are evident in the project's ability to deliver more accurate and precise estimates, foster data-driven decision-making, enable real-time monitoring and

adaptation, enhance predictive capabilities, reduce costs, offer customization and flexibility, and facilitate scenario analysis for risk mitigation. These advantages collectively contribute to the project's potential to significantly enhance project management in a business context.

However, we have also acknowledged the potential pitfalls, including data quality challenges, complexity in implementation, data privacy and security concerns, resource intensiveness, integration challenges, and the need for initial setup and learning. These disadvantages underscore the importance of careful planning, investment in training and infrastructure, and robust data governance to ensure the success of the project.

In conclusion, the "Estimation of Business Projects Using Data Analytics" project represents a transformative journey that demands a well-balanced approach to maximize its benefits while addressing its challenges. As we move forward, it is imperative to focus on data quality, security, and the development of human resources to fully realize the potential of data analytics in the estimation of business projects. With the right strategies and a commitment to data-driven decision-making, this project has the potential to reshape the landscape of project management, ultimately driving business success and growth.

12.FUTURE SCOPE

The future scope for the project "Estimation of Business Projects Using Data Analytics" holds immense potential for innovation and transformation. As technology evolves, the project can explore advanced machine learning and AI integration for even more accurate and insightful predictions. It may expand into Agile project management and consider IoT and sensor data for real-time adjustments to project plans. The incorporation of blockchain can enhance data security and transparency, while seamless integration with project management software can improve user adoption. Cross-industry applications and a focus on ethical AI and responsible data use are on the horizon, and as quantum computing matures, it can be harnessed for complex modeling. The development of data-driven decision support systems, global standards, and interdisciplinary collaboration are also key elements of the future scope. With a commitment to continuous improvement and adaptation to emerging trends, this project is poised to lead the way in data-driven project estimation, offering benefits that extend across various sectors and industries.

APPENDIX

SOURCE CODE

Github

<https://github.com/Mohamedfawaz786/Naan-Mudhalvan/blob/main/Final%20D>