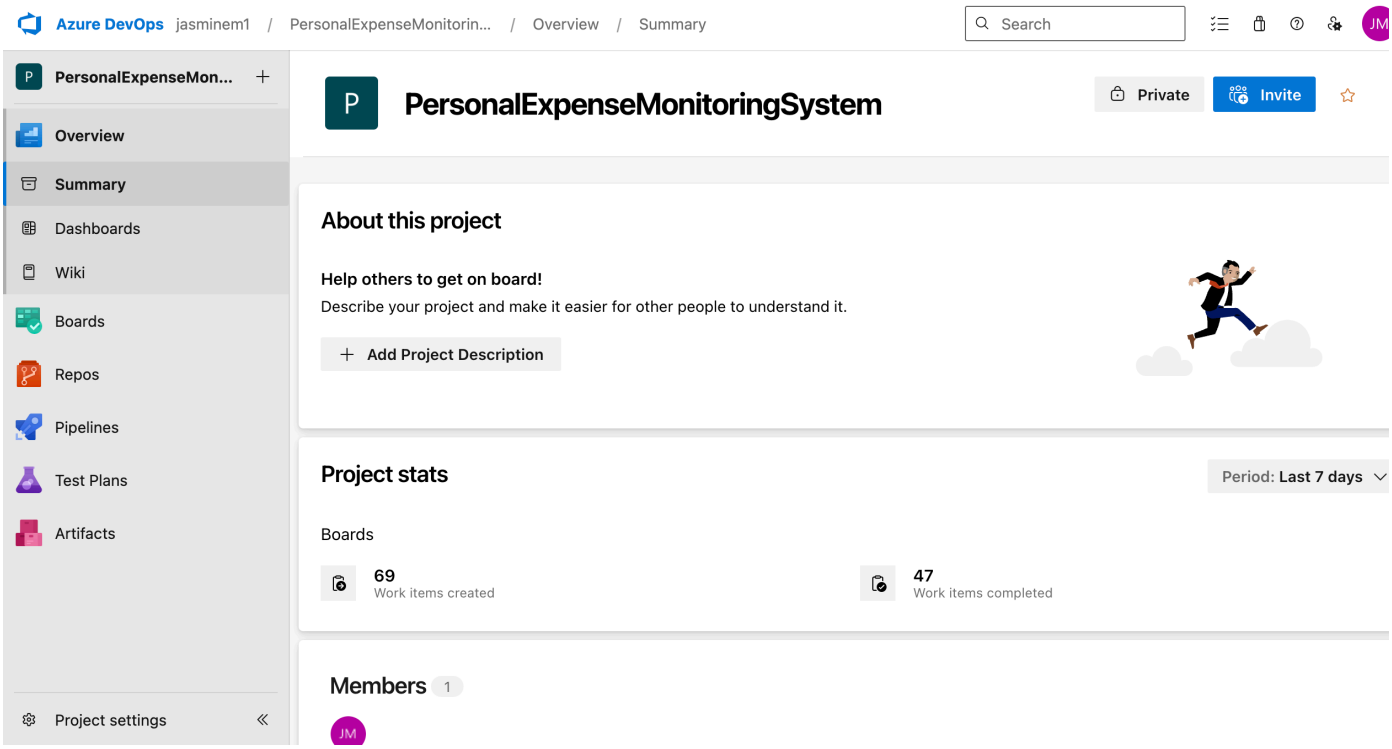
OVERVIEW FOR PERSONAL EXPENSE MONITORING

Objective:

Build a basic system to track and analyze personal or household expenses. The system should help users categorize spending, identify unusual patterns, and automate monthly summaries.



Epic

Personal Expense Monitoring System

Features

- Expense Data Storage
- Category Management
- Receipt Upload & Storage
- Expense Analysis & Summaries
- Anomaly/Spike Detection
- ETL & Data Integration

- Automated Reporting

User Stories

Expense Data Storage

- MySQL User Table
- MySQL Expense Table
- MySQL Category Table

Category Management

- Add Category
- Edit Category
- Delete Category

Receipt Upload & Storage

- Store Receipt Data (MongoDB)
- Upload & Retrieve Scanned Receipts

Expense Analysis & Summaries

- Monthly Expense Summaries
- Category-Wise Expense Breakdown

Anomaly/Spike Detection

- Unusual Expense Alert (PySpark)
- Detect Large One-Time Transactions

ETL & Data Integration

- Combine User & Expense Data (Databricks)
- Create Summary/Cleaned Tables

Automated Reporting

- Schedule Analysis Pipeline (Azure DevOps)
- Automated CSV/Delta Report Generation

Tasks

MySQL User Table

- Create user table schema in MySQL
- Add user CRUD operations

MySQL Expense Table

- Create expense table schema
- Add expense CRUD operations
- Write stored procedure for monthly category totals

Category Management

- Implement category CRUD in MySQL

Receipt Data (MongoDB)

- Design JSON schema for receipts/notes
- Insert sample receipt in MongoDB
- Add indexes for user and receipt lookup

Data Cleaning (Python)

- Write Python script to import and clean data from CSV/API
- Standardize date and amount formats

Expense Summaries

- Use pandas to create category-wise breakdown
- Compute monthly totals and averages with numpy

Anomaly Detection (PySpark)

- Load large expense dataset in Spark
- Group by user and month, detect spikes/anomalies
- Output flagged users/expenses

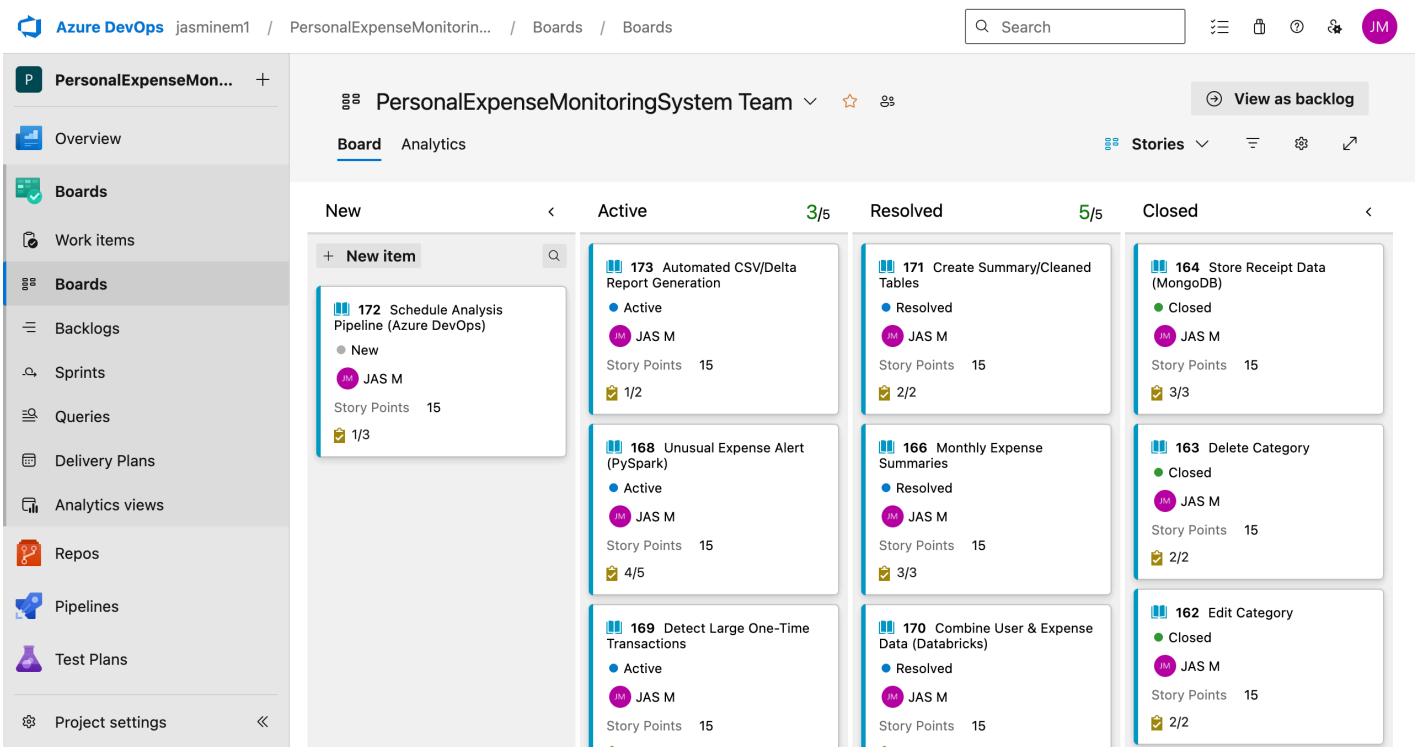
Data Integration & ETL (Databricks)

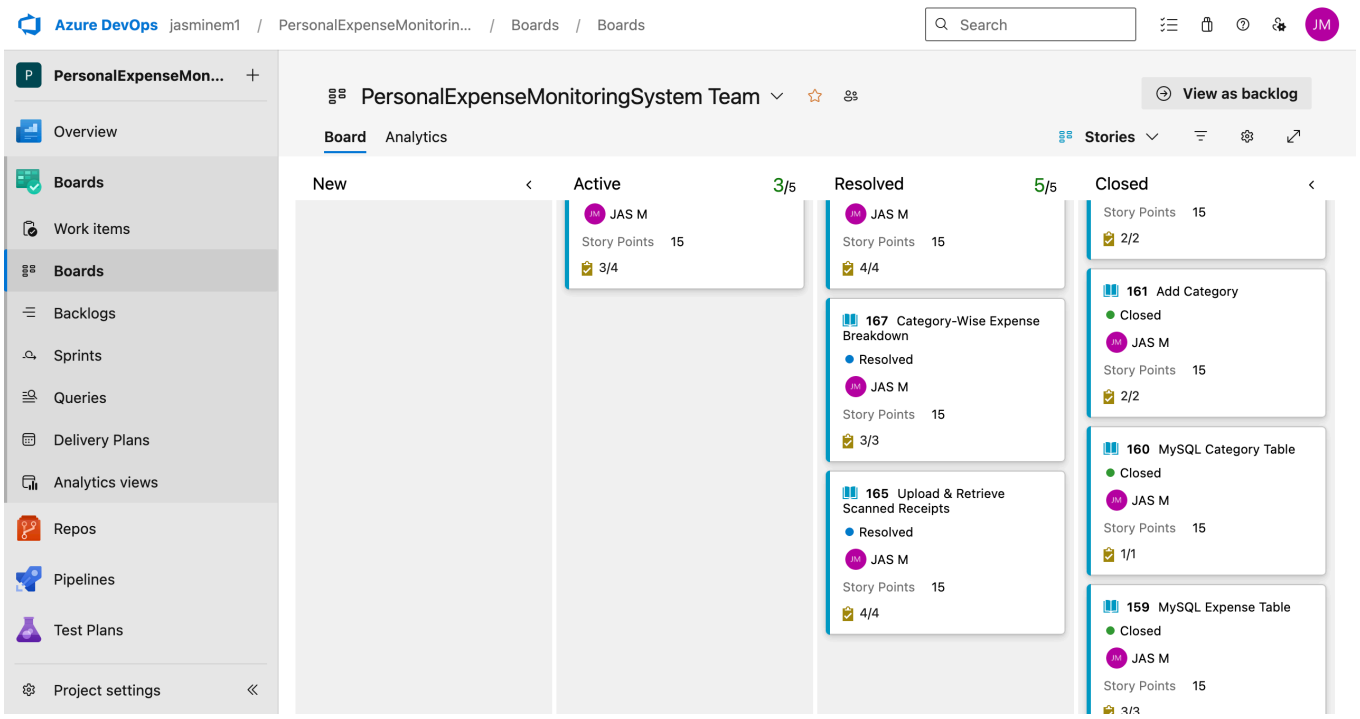
- Upload cleaned data to Databricks
- Join user and expense tables
- Generate summary table for dashboard
- Export results as Delta/CSV

Automated Pipeline (Azure DevOps)

- Build YAML pipeline for weekly/monthly execution
- Output monthly CSV report
- Configure savings alert if spending exceeds threshold

BOARDS





Summary:

This project is a comprehensive personal or household finance tracker designed to help users record, categorize, and analyze their expenses. The key objectives are to provide clear visibility into spending patterns, automate reporting, and detect any unusual spending activity. The system uses MySQL for structured data (users, expenses, and categories), MongoDB for unstructured data (like receipt notes and scans), and Python with Pandas/NumPy for data processing and analysis. For large datasets and automation, PySpark and Azure Databricks are used for ETL, and an Azure DevOps pipeline runs analysis and reporting automatically.

Major Features:

- Secure storage of expenses, users, and category data
- CRUD operations for easy management of expenses and categories
- Upload and retrieval of digital receipts and notes
- Automated monthly summaries and category-wise breakdowns
- Anomaly detection for spikes or unusual transactions
- Seamless data integration using Databricks and automated pipeline
- Output of insights as reports and savings alerts, delivered on a schedule

This system delivers end-to-end automation, actionable insights, and robust data handling for effective personal finance management.