Financial Receipt Capture and Analysis System

Team members: Will Southerland, Josh Tarongoy, Lindsey Marandina, & Josh Whitchurch| **Faculty adviser:** Thang Dinh, Ph.D. | **Sponsor:** Capital One | **Mentor:** Mehul Garnara

Problem Statement

Many people face challenges keeping track of spending, with many often spending more than what they can afford leading to credit card bills racking up. In our digital world, the manual processes of storing, searching, and analyzing physical receipts is often:

- time-consuming
- Inefficient
- prone to errors
- Disorganized
- difficult to analyze

making it challenging for individuals to track their spending and make informed financial decisions. Many customers are unable to make well informed financial decisions beyond individual transactions within each financial institution.

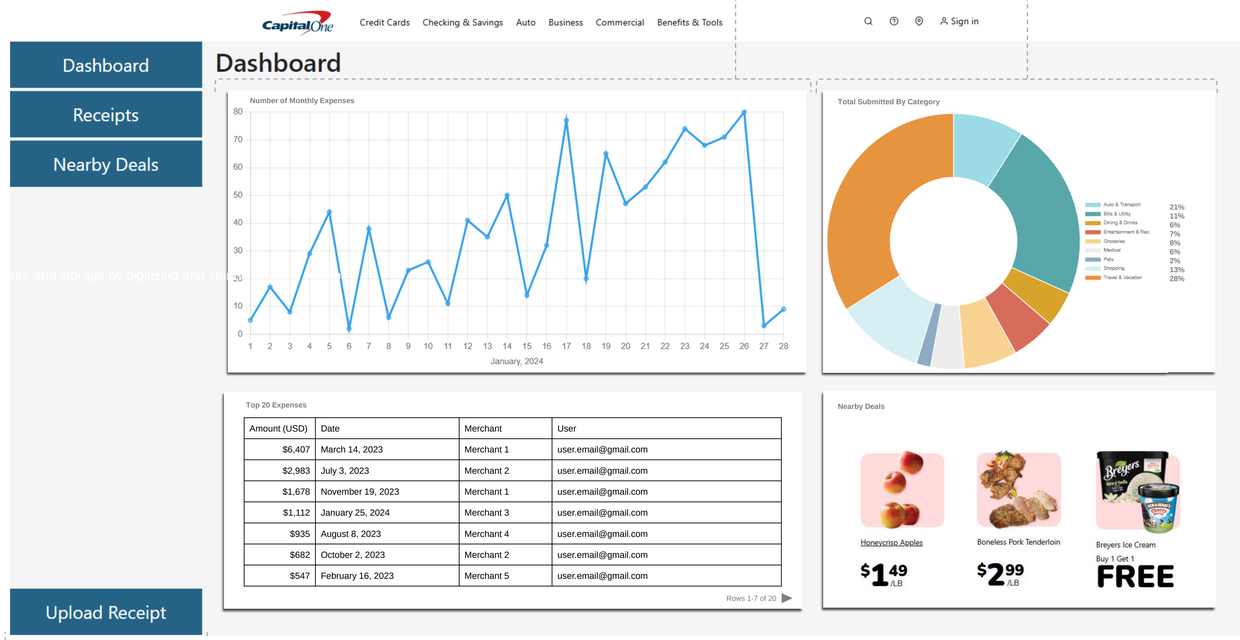
Project Objective

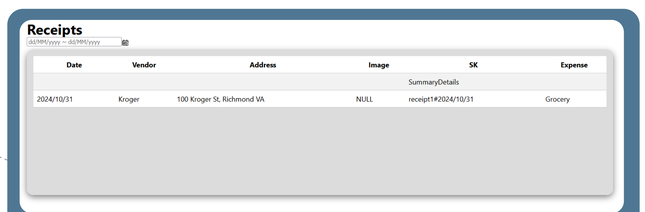
This project aims to develop an application that streamlines the process of capturing and managing hard copy credit card transaction receipts through capture, storage, and analysis of physical receipts. The system will allow users to:

- Search receipts by date, vendor, category, and other criteria.
- Analyze spending across different periods and categories.
- Identify spending trends to gain valuable insights.
- Utilize a categorization system to enhance financial tracking and analysis.

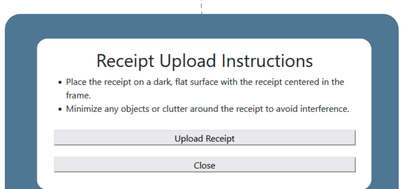
Spending analysis through comparing and analyzing spending over time and across spending categories. Trend Identification through identifying spending patterns to inform financial decisions.

Categorization by automatically and/or manually classifying receipts based on common spending categories.





Search functionality that allows users to locate specific receipts efficiently.



Receipt capture and storage by digitizing and storing physical receipts.

Our Approach

- Users upload a preexisting photo of a receipt or capture it with their mobile device.
- The image is preprocessed to ensure readability by the OCR (Optical Character Recognition).
- Tesseract extracts the total, date, and vendor from the image. If Tesseract fails, the image is sent to AWS Textract for extraction. If both OCRs fail, the user is prompted to manually enter the data.
- The data is passed to DynamoDB to store the receipt data.
- The data is visualized through donut charts and graphs to easily view spending habits.
- Data can be filtered by date range, expense type, total spent, etc.

spaCy pandas JS aws

VCU College of Engineering