Tech Stack Overview and Interactions

1. **React**:
   - **Role**: Frontend for a web-based application.
   - **Interactions**:
     - Renders visualizations of processed data.
     - Sends user inputs and requests (e.g., upload receipts, fetch analytics) to the Node.js backend via HTTP POST/GET requests.
2. **Node.js**:
   - **Role**: Backend server that acts as an intermediary between the React frontend and AWS services.
   - **Interactions**:
     - Receives API requests from the React app.
     - Processes requests (e.g., validating data, managing uploads) and communicates with AWS services (Lambda, DynamoDB, S3).
     - Returns responses to the React frontend with data fetched or processed from AWS.
3. **AWS Lambda**:
   - **Role**: Serverless compute service for executing code in response to events.
   - **Interactions**:
     - Validates incoming data and performs OCR on uploaded receipt images using Textract.
     - Triggers on S3 upload events, processes data, and stores results in DynamoDB.
4. **Amazon S3**:
   - **Role**: Object storage for unstructured data.
   - **Interactions**:
     - Stores raw receipt images and metadata to be processed by Textract.
     - Holds processed data lakes where analysis outputs or transformed datasets can be stored for further processing.
5. **Amazon Textract**:
   - **Role**: OCR service for extracting text and structured data from receipt images.
   - **Interactions**:
     - Receives images from S3 and returns extracted text and structured data to AWS Lambda, which then forwards this data to DynamoDB.
6. **DynamoDB**:
   - **Role**: NoSQL database for storing structured data.
   - **Interactions**:
     - Stores user information and receipt metadata extracted from Textract.
     - Provides fast retrieval for the Node.js backend to respond to frontend requests.
7. **AWS Glue**:
   - **Role**: ETL (Extract, Transform, Load) service for preparing data.

- **Interactions**:
  - Extracts metadata from DynamoDB (or S3) for transformation and analysis.
  - Cleans and prepares data for further analysis,

8. **Amazon CloudWatch**:
   - **Role**: Monitoring and observability service.
   - **Interactions**:
     - Monitors logs and metrics from the backend services to ensure smooth operation.
     - Tracks data flow between DynamoDB and S3, alerting for any data loss or processing issues.

9. **Amazon EC2**:
   - **Role**: Virtual server for model training.
   - **Interactions**:
     - Hosts machine learning models and training scripts, pulling data from S3 for training.
     - Periodically retrains models using updated data to improve accuracy and adapt to new trends

10. **Black Box (Numpy, Pandas, PyTorch)**:
    - **Role**: Analysis of processed data.
    - **Interactions**:
      - Analyzes structured data stored in S3 to identify trends and insights.
      - Generates reports or visualizations based on the analysis.

Frontend

React

Backend

Node.js

user inputs/
requests

results for UI

API request

returned data

AWS Services

| Lamda | S3 | EC2 |

Textract (OCR)

Cloudwatch

DynamoDB

Glue