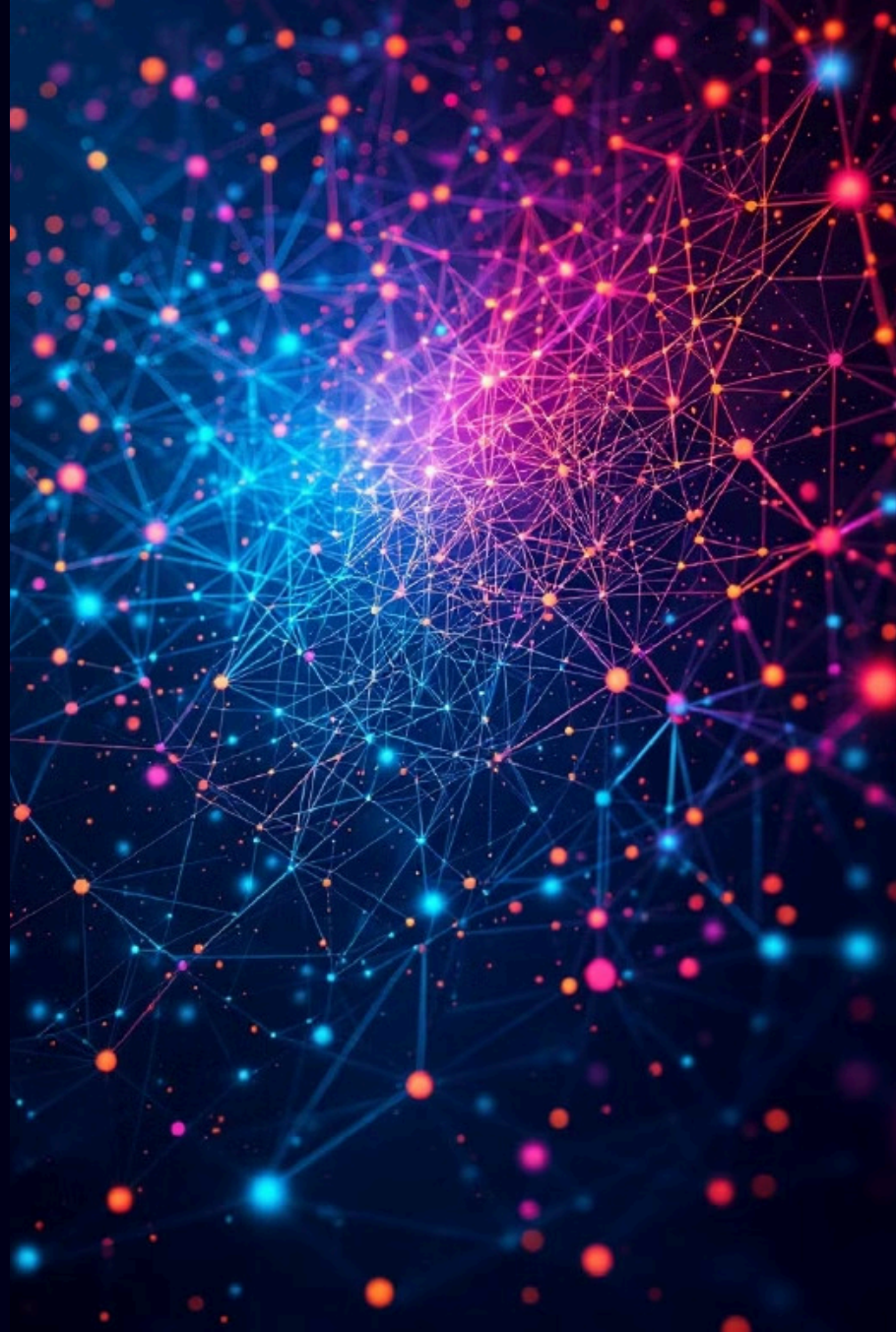
HR Data Risk Modeling & Scenario Analysis

A fully serverless, event-driven automation system for HR data analysis. Built with AWS Lambda, S3, and SNS, it processes data through clear stages and delivers interactive dashboards.



Cloud Automation Architecture

1

Data Fetching

External data retrieval from Google Drive/APIs via EventBridge triggers.

2

Processing

S3 event notifications trigger polynomial regression analysis.

3

Visualization

Dashboard generation automatically triggered by processed data events.

Data Processing Pipeline

Fetch Data

Lambda function fetch_0_data retrieves external data and loads to S3.

Polynomial Regression

poly_regression analyzes data with confidence intervals and risk scenarios.

Dashboard Creation

dashboard function generates interactive visualizations with Plotly.

Infrastructure as Code

S3 Storage

Organized bucket structure with dedicated folders for raw, processed, and archived data.

IAM Policies

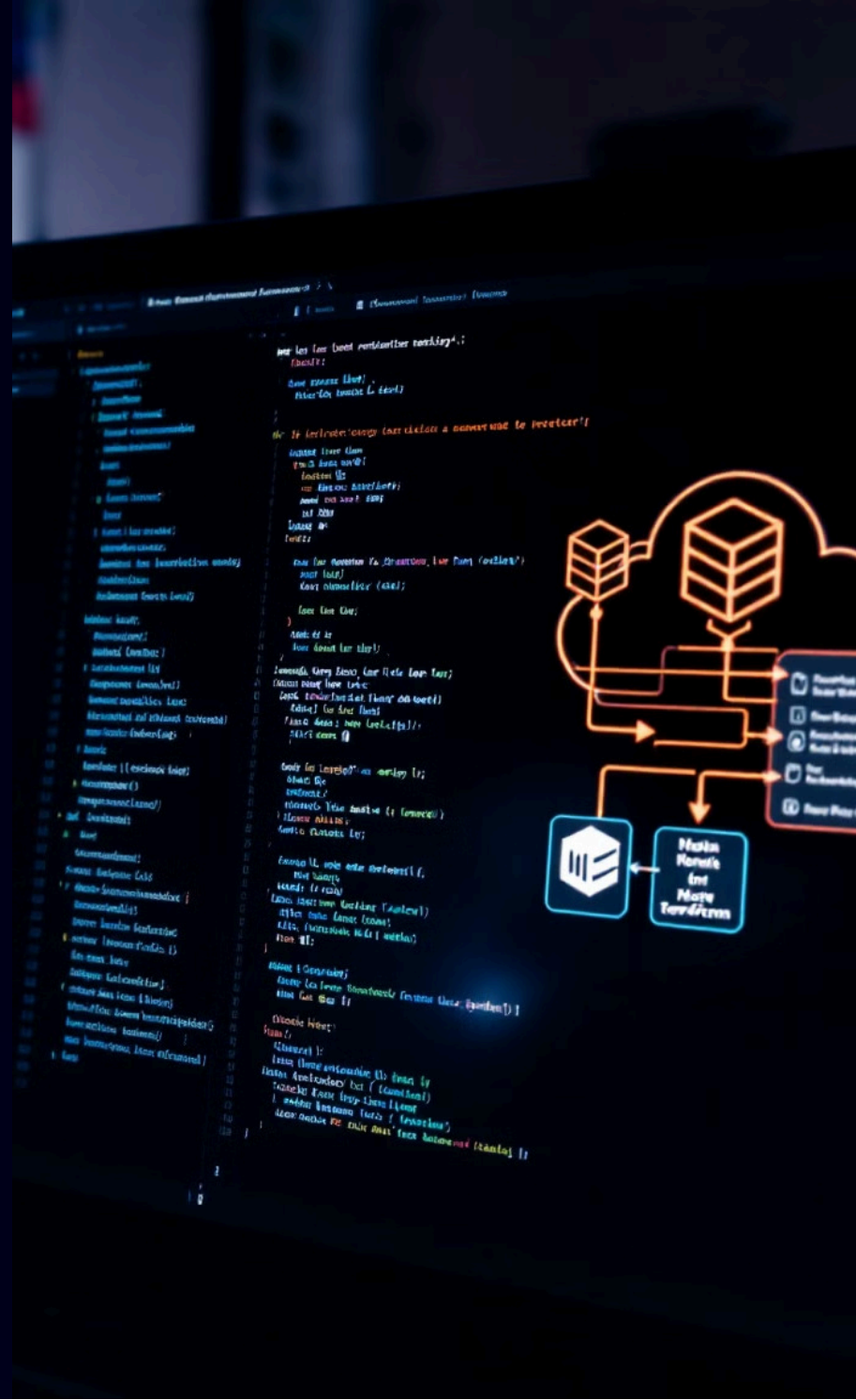
Secure Lambda access with precisely defined permissions for S3 operations.

Lambda Functions

Three specialized functions defined with appropriate runtime and handlers.

SNS Notifications

Real-time alerts through hr_data_notification topic for monitoring.



Containerized Environment

HR Data Repository

Docker container with Python 3.11 Lambda base image. Includes specialized libraries for polynomial regression analysis.

Dashboard Repository

Separate container optimized for visualization. Contains Plotly and other dashboard generation dependencies.

Benefits

Consistent environments across development and production. Simplified deployment and scaling capabilities.

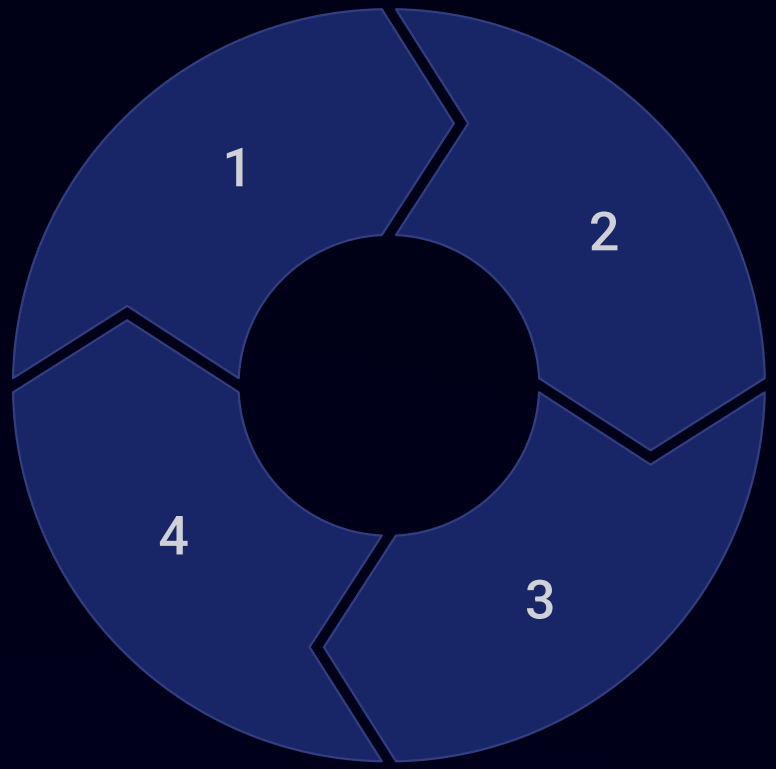
Advanced Data Analysis

Polynomial Regression

Mathematical modeling of non-linear HR data trends.

Actionable Insights

Data-driven recommendations for HR decision making.



Confidence Intervals

Statistical boundaries for prediction accuracy.

Risk Scenarios

Multiple outcome projections based on variable conditions.

Interactive Dashboards

Plotly-powered visualizations accessible at: https://s3.eu-north-1.amazonaws.com/hr.data/Dashboard/index.html provide scenario analysis for improved data interpretation.

Technical Stack & Libraries

AWS Services

- Lambda for serverless computing
- S3 for data storage
- SNS for real-time notifications

Python Libraries

- Pandas for data manipulation
- Plotly for interactive visualizations
- Scikit-learn and Statsmodels for analysis

DevOps Tools

- Docker for containerization
- Terraform for infrastructure
- ECR for container registry