# Football FFP Analysis - System Design

## 1. Overview

The Football FFP (Financial Fair Play) project is a data pipeline designed to scrape, store, analyze, and visualize football club financial data in compliance with FFP regulations. The project leverages AWS services and Python for data processing, storage, analysis using Bedrock, and visualization via QuickSight.

## 2. Architecture Overview

The architecture involves five main components:  
1. Scraper: Collects financial data from football club sources.  
2. S3: Stores raw and processed data.  
3. Bedrock: Analyzes data using foundation models.  
4. OpenSearch: Indexes processed insights for querying.  
5. QuickSight: Visualizes the processed data interactively.

## 3. System Components

### 3.1 Data Ingestion (Scraper)

The Python scraper collects financial data from various football clubs. The collected data is saved locally before being uploaded to AWS S3.

### 3.2 Storage (S3)

The scraped data is uploaded to an S3 bucket for persistent storage and further processing.

### 3.3 Data Analysis (Bedrock)

The uploaded data is sent to an AWS Bedrock model which generates financial insights based on club data. The results are stored in OpenSearch for query and retrieval.

### 3.4 Indexing & Search (OpenSearch)

Insights and enriched metadata are indexed in OpenSearch for semantic querying and visualization.

### 3.5 Visualization (QuickSight)

Data in S3 and OpenSearch is used to generate visual dashboards using Amazon QuickSight, allowing stakeholders to interactively explore financial compliance and insights.

## 4. Deployment & Execution

Deployment is handled via a Bash script (`deploy.sh`) which sets up necessary AWS infrastructure using CloudFormation. Scripts are run sequentially to perform scraping, uploading, analysis, and visualization setup.