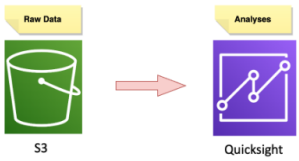
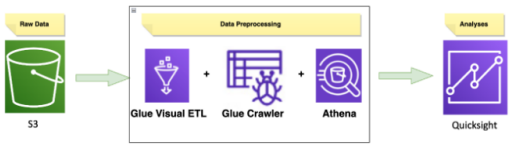
Project1: Cloud Foundations Project Narrative Version

- **Purpose**
  - This project is dedicated to deepening CSA Interns' understanding of various AWS services. Interns will utilize at least two services to create a detailed presentation comparing their functionalities.
- **Background and Problems**
  - A company is seeking solutions to maintain road infrastructure as more people switch to electric vehicles. they have a dataset in CSV format that needs to be **stored in the cloud** and **retrieved immediately**. Although **data loss is acceptable**, a backup storage solution is still required. The company is requesting assistance in selecting the appropriate AWS services to effectively implement the road charge system and **visualize** their data.
- **Recommendations (demo video)**
  - **Option 1:**
    - 

    - Based on the requirement, we recommend using QuickSight to read data directly from an S3 bucket. Firstly, S3 is an object storages service with multiple storage classes, allowing us to select the appropriate class for different situations as the business grows. S3 is region-based, which ensures data durability and availability by storing it across multiple, isolated locations within the chosen region. QuickSight is a serverless service that can integrate directly with S3. Its pay-per-session pricing can reduce BI solution costs by up to 74% over three years.
      - **Pros:**
        - Using S3 and QuickSight allows for quick setup, direct access to data, and minimal configuration.
      - **Cons:**
        - Less flexibility in data preprocessing and a direct dependency on the format and structure of data in S3
  - **Option 2:**
    - 

    - If the format and structure of the data are problematic, we can use Glue and Athena for preprocessing. First, we can use AWS Glue Visual ETL to extract data from sources, transform it into a usable format, and load it into an S3 bucket for analysis, reporting, or further processing. AWS Glue crawlers can then automatically organize the data, making it easy to analyze with Athena. Note: need to set up permission policies for accessing Athena, S3, and Glue.
      - **Pros**:
        - Using Glue and Athena provides powerful data processing and transformation capabilities, as well as flexible data querying with scalable SQL in Athena.
      - **Cons:**
        - More setup requirements and potentially higher costs due to the use of multiple services.

- **Next step (demo video)**
    - Use S3 replication rules to implement backup storage. Based on the requirements, we choose One Zone-IA, a storage class in Amazon S3 designed for infrequently accessed data that still requires rapid access when needed. Since data loss is acceptable, S3 One Zone-IA stores data in a single availability zone and costs 20% less than S3 Standard-IA.
- **FAQs**