**Architecture Design**

**Amazon Sales Data**

**Analysis**

Revision Number: 0.1

Deepraj Arya

**DOCUMENT CONTROL**

**Change Record:**

|  |  |  |  |
| --- | --- | --- | --- |
| **VERSION** | **DATE** | **AUTHOR** | **COMMENTS** |
| 0.1 | 12- Oct -  2023 | Deepraj Arya | Architecture Design |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**Reviews:**

|  |  |  |  |
| --- | --- | --- | --- |
| **VERSION** | **DATE** | **REVIEWER** | **COMMENTS** |
| 0.1 | 25-Oct-2023 |  |  |

**Approval Status:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **VERSION** | **REVIEW**  **DATE** | **REVIEWED BY** | **APPROVED BY** | **COMMENTS** |
|  |  |  |  |  |

**Architecture Design**

1. Data Ingestion and Storage:
   1. **Data Sources:** Data may come from various sources like sales databases, e-commerce platforms, or external sources. Ingest data from these sources. This is project of PW so they provided the data for this project.
   2. **Data Storage:** Store the data in a database or data warehouse. Consider using technologies like PostgreSQL, MySQL, or cloud-based solutions like Amazon Redshift or Google BigQuery for scalability.
2. **Data Preprocessing:**
   1. **ETL (Extract, Transform, Load):** Implement ETL processes to clean and preprocess the data. This includes handling missing values, outliers, and transforming the data into a usable format.
3. **Exploratory Data Analysis (EDA):** Use tools like Pandas, Matplotlib, or Seaborn to perform EDA. Visualize historical sales data to understand patterns, seasonality, and correlations.
4. **Feature Engineering:** Create features that can impact sales trends. This includes generating lag features, rolling statistics, and integrating external data (e.g., holidays, promotions).
5. **Time Series Data Handling:** Prepare time series data for modelling. This includes dealing with seasonality and stationarity.