**Epsilon project for data engineer intern**

**About the company:-**

Link- <https://en.wikipedia.org/wiki/Epsilon>

Epsilon is a global advertising and marketing technology company positioned at the center of Publicis Groupe. We connect advertisers with consumers to drive performance while respecting and protecting consumer privacy and client data. Epsilon accelerates clients’ ability to harness the power of their first-party data in order to enhance, activate and measure campaigns with confidence. We believe in an open, privacy-first advertising ecosystem. Over decades, we’ve built the industry’s most comprehensive identity graph to give brands, agencies, and publishers the ability to reach real consumers across all channels and the open web.

**Link-**[**https://www.epsilon.com/us/about-us/who-we-are**](https://www.epsilon.com/us/about-us/who-we-are)

**Project Title:**

**Developing a Data Quality Framework for Enhanced Decision-making**

**Project Overview**

Epsilon is a global marketing technology company that helps businesses acquire, retain, and grow their customers. In order to make better decisions about their marketing campaigns, Epsilon needs to ensure that the data they use is of high quality.

This project will develop a data quality framework for Epsilon. The framework will define the standards for data quality, identify the root causes of data quality problems, and develop solutions to improve data quality.

**Tasks**

**The following are some of the tasks that a data engineer intern could work on for this project:**

* Conduct a data assessment to determine the current state of data quality
* Develop a requirements document that outlines the functionality and performance requirements for the data quality framework.
* Design and implement the data quality framework.
* Develop a data quality improvement plan to address the root causes of data quality problems.
* Test and deploy the data quality framework.
* Provide training and support to Epsilon's users so that they can use the data quality framework to improve the quality of their data.

**The project will be completed in three phases:**

**Phase 1:** In this phase, the project team will conduct a data assessment to determine the current state of data quality at Epsilon. The team will also develop a requirements document that outlines the functionality and performance requirements for the data quality framework.

**Phase 2:** In this phase, the project team will design and implement the data quality framework. The team will also develop a data quality improvement plan to address the root causes of data quality problems.

**Phase 3:** In this phase, the project team will test and deploy the data quality framework. The team will also provide training and support to Epsilon's users so that they can use the data quality framework to improve the quality of their data.

**Benefits**

**The benefits of this project include:**

* Improved decision-making: Epsilon will be able to make better decisions about their marketing campaigns by leveraging the insights gained from their data.
* Increased customer satisfaction: Epsilon will be able to better understand their customers and their needs, which will lead to increased customer satisfaction.
* Reduced costs: Epsilon will be able to reduce costs by optimizing their marketing campaigns and eliminating waste.

**Risks**

**The risks associated with this project include:**

* Data quality: The quality of the data at Epsilon is critical to the success of this project. If the data is not of high quality, the data quality framework will not be effective.
* Technical challenges: The development and implementation of a data quality framework is a complex technical challenge. There is a risk that the project team will not be able to overcome these challenges.
* Change management: The implementation of a data quality framework will require changes to Epsilon's business processes and culture. There is a risk that these changes will not be successful.

**Mitigation Strategies**

The project team will mitigate the risks associated with this project by taking the following steps:

* **Data quality**: The project team will work with Epsilon's data stewards to ensure that the data at Epsilon is of high quality.
* **Technical challenges**: The project team will bring in experts with experience in data quality framework development and implementation.
* **Change management:** The project team will work with Epsilon's business users to develop a change management plan.

**Conclusion**

As the Data Engineer Intern, the completion of the "Developing a Data Quality Framework for Enhanced Decision-making at Epsilon" project marks a significant milestone in your journey to contribute to Epsilon's data-driven decision-making processes. The project has successfully delivered a robust and comprehensive data quality framework, enabling Epsilon to ensure the integrity, accuracy, and reliability of data used for decision-making.

Throughout the project, you have demonstrated your skills and commitment in data profiling, validation, cleansing, and monitoring. Your contributions have been vital in establishing a data quality framework that addresses the existing challenges and improves the overall data quality within Epsilon's data infrastructure.

The project's objectives have been achieved successfully. The assessment of current data quality challenges provided valuable insights into the areas of improvement. By defining data quality metrics and benchmarks, you have established a baseline to measure and track data quality improvement over time. The data profiling and assessment activities have identified data anomalies and issues, forming the foundation for subsequent data quality enhancement efforts.

Through the implementation of data validation rules and automated data cleansing processes, you have laid the groundwork for improving data integrity and accuracy. The data quality monitoring framework developed during the project will enable Epsilon to proactively identify and address data quality issues, ensuring continuous improvement in data quality standards.

The project has also contributed to enhancing decision-making processes at Epsilon. The availability of data quality reports and dashboards provides visibility into data quality metrics and trends, empowering stakeholders to make informed decisions based on high-quality data. The collaboration with business stakeholders and IT teams has ensured that the data quality framework aligns with business objectives and requirements.

As a Data Engineer Intern, you have played a vital role in the project's success. Your dedication, collaboration, and commitment to delivering high-quality results have been commendable. The documentation of the data quality framework, processes, and guidelines will serve as a valuable resource for future reference and maintenance.

Moving forward, it is essential to ensure smooth knowledge transfer and handover of the project outcomes to the relevant teams. By sharing your project insights, lessons learned, and recommendations, you can contribute to the continued improvement of data quality practices at Epsilon.

The "Developing a Data Quality Framework for Enhanced Decision-making at Epsilon" project has provided you with valuable hands-on experience in data engineering and data quality management. It has equipped you with the skills and knowledge required to contribute effectively to data-driven initiatives in your future career endeavors.

Overall, the successful completion of this project has laid the foundation for improved data quality and decision-making processes at Epsilon. Your contributions have made a positive impact, and the project sets the stage for further advancements in data quality management within the organization.