## Weeks 11 & 12: Final Project Assignment

### Objective:

Leverage the technologies you’ve learned over the course of this class to implement a data-focused solution of your choosing. At the end of this project, you should create a report documenting your work.

### Requirements:

1. You must use at least 2 of the components you’ve learned (e.g., HDFS, YARN, Hive, HBase, Spark, Kafka, Solr, NiFi).
2. Document your data source.
3. Describe the operations or transformations you performed on the data.
4. Explain why you chose to use the components you did.
5. Include screenshots of the data usage across all the components you used.
6. Your report should be at least 2 pages in length.

### Suggested Use Cases (Optional):

While you are free to choose any data and scenario for this project, here are some sample use cases and associated datasets to inspire your work:

1. **E-Commerce - Product Inventory Management:**
   * **Problem:** Ingest sales data and manage product inventory in real-time, offering insights into stock levels, sales trends, and reorder points.
   * **Sample Data:** [E-Commerce Dataset](https://www.kaggle.com/carrie1/ecommerce-data)
2. **Finance - Transaction Log Analysis:**
   * **Problem:** Streamline large volumes of transaction logs, identifying transaction volumes, peak transaction times, and other important metrics.
   * **Sample Data:** [Bank Transactions Dataset](https://www.kaggle.com/ntnu-testimon/banksim1)
3. **Social Media - User Interaction Analysis:**
   * **Problem:** Ingest and analyze user interactions on a social media platform, deriving insights about user engagement, most active times, and trending posts or topics.
   * **Sample Data:** [Social Media Interaction Dataset](https://www.kaggle.com/c/predicting-user-engagement-data)
4. **Tourism - Hotel Booking Analysis:**
   * **Problem:** Process and analyze hotel booking data to identify trends, booking preferences, and cancellations.
   * **Sample Data:** [Hotel Booking Dataset](https://www.kaggle.com/jessemostipak/hotel-booking-demand)

Remember, the use of these suggested datasets and use cases is optional. You are encouraged to explore other datasets that may be of interest to you.

### Submission Guidelines:

1. Submit your report in PDF format.
2. Ensure that all images and text in your report are legible.
3. Adhere to the 2-page minimum requirement.
4. Ensure to provide a link or reference to any external datasets or sources you’ve used.

### Evaluation:

Your final project will be evaluated based on: 1. The complexity and relevance of your chosen use case. 2. How effectively you’ve leveraged the technologies learned in class. 3. The clarity and organization of your report.

Good luck, and we’re excited to see what you come up with!