**Week One Status**

**9/14/2022**

**TEAM RAD**

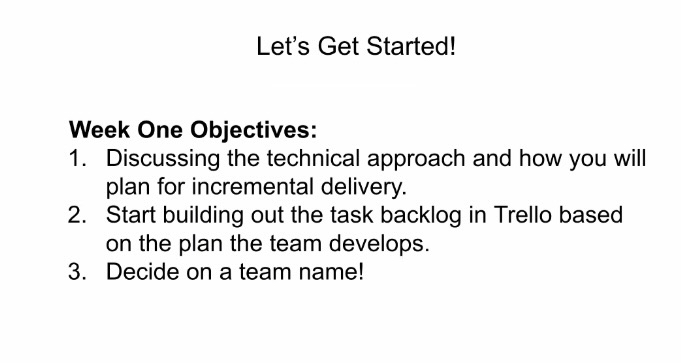
**Real-time Anomaly Detection**

**Gavin Cutchin -**

**Gabriella Rivera -**

**Jared Dunn -**

**Questions: Is this only going to be deployed in AWS? Can we use AWS proprietary technologies?**



**Parsing and Baselining**

What language do we want to write this in? **Python**

Would data science libraries like Pandas or Spark be useful here? **Spark (PySpark)**

What kind of model is best to capture a baseline of activity? **Using AWS sageMaker to suggest models based on baseline data**

What types of anomalies in activity should we consider (volume, features of

the events themselves, etc)? (**Part of Backlog)**

How could we approach this “cloud first” to make it scalable and easily

deployable in AWS? **Use AWS technologies.**

**Streaming Process to Detect Anomalies**

What sort of streaming data technologies should we employ? **AWS Kinesis**

How can we capture anomalies taking place over a longer period of time while

still consuming the data in real time? **Hard coded case in SageMaker (needs research)**

How should we store the results of anomaly detection to later display in the

web application? **AWS S3 Bucket**

How can we capture anomalous relationships between events in real time? **AWS Athena for real time analytics**

How could we approach this “cloud first” to make it scalable and easily

deployable in AWS? **Do everything in AWS**

**Display in Web App**

How can we display the anomaly information in a way that is useful to a threat

Analyst? **Kibana with AWS Elasticsearch**

How could we approach this “cloud first” to make it scalable and easily

deployable in AWS? **Build using AWS libraries**