

Hamza Farhat Ali

100657374

Comparisons

- Differences:
 - DataPrep and DataFlow clusters are created using automation while DataProc relies on manual setup for a cluster.
- Advantages:
 - Dataflow has many use cases such as real time AI, stream analytics and sensor data processing for real time scenarios.
 - High security as the dataset is protected due to the access management for users.
- Disadvantage:
 - The support decreases as GCP has limited data centers across the globe which leads to the limitations listed below.
- Limitations:
 - GCP DataFlow has a ceiling on the maximum stream value at one time.
 - GCP DataFlow has a ceiling on custom job metrics to 100.
 - GCP DataFlow only allows for 15 000 messages every 30 seconds.