**Questions:**

**What are the risks involved in building such a pipeline?**

**Ans:** Streaming API can disconnect, and we will not receive any tweets data.

As the data gets larger, scaling pipelines becomes a big problem. For starters, backups take longer and eat up a bunch of resources. Rebuilding indexes, defragmenting storage and reorganizing data are all time- and resource-consuming operations

**How would you roll out the pipeline going from proof-of-concept to a production-ready solution?**

**Ans:** Once I am done with the Proof of Concept, I will evaluate the outcome of my pipeline against the original goals and assumptions. If my POC met earlier assumptions or even exceeded them, it means I am on the right track.

I will then test pipeline with large-scale data and in different environments.

**What would a production-ready solution entail that a POC wouldn't?**

**Ans:** Program reusability, Data storage and monitoring, Governance, Scalability

**What is the level of effort required to deliver each phase of the solution?**

**Ans:** 2 weeks for each phase

**What is your estimated timeline for delivery for a production-ready solution?**

**Ans:** 4-6 weeks (Assuming starting from proof of concept)

POC, Development, Testing, Deployment, and postproduction support – 2 weeks for each phase