**Data Analyst Challenges**

* Query taking too long
* No easy way to combine data
* Unsustainable servers
* Clusters avert scaling
* High cost
* No central data store

**Why cloud is good**

* Storage is cheap
* Focus on queries and infrastructure
* Massive Scalability
* **Hard Drives**
* **Processing Power/Servers**
* **Network**
* **Database Technicians**

**Data Analyst Tasks**

* Ingest – Take data in
* Transform – Prepare, clean, and transform data
* Store – Create, save, and store data sets
* Analyze – Derive insights from data
* Visualize – Explore and present data insights

**Data Analysis Challenges**

* Ingest – Petabytes of data is going to bottleneck the tool, Data volume, Data variety, Data velocity
* Transform – Slow exploration, Slow processing, Unclear logic
* Store – Storage cost, Hard to scale, Latency issues
* Analyze – Slow queries, Data volume, Sliced data
* Visualize – Dataset size, Tool latency
* Don’t manage infrastructure yourself
* Focus on finding insights

**9 Fundamental BigQuery Features**

* Fully-managed data warehouse – No-ops petabyte-scale
* Reliable - Backed by Google data centers
* Economical – Pay only for the processing and storage you see
* Secure – Role ACLs, data encrypted in transport and at rest
* Auditable – Every transaction logged and query able
* Scalable – Highly parallel processing model means fast queries

**SQL + WEB UI**

* Flexible, fast, and familiar

**Data Preparation Tools**

* GUI for exploring columns and rows
* Fast summary statistics

**Visualization Tools**

* Visually shape and reshape quickly
* See data a different way

**Steps to explore data through SQL**

* Ask good questions
* Know your data
* Write good SQL

**BigQuery Jobs**

* Tasks given a unique id by the web ui
* Can run concurrently
* History stored for 6 months

**Four Types of BigQuery Jobs**

* Query Data
* Load Data
* Extract Data
* Copy Data