



# Opening Remarks

**Evren Eryurek, PhD**

Director, Google Cloud, Data Analytics

@evren\_eryurek



# Why Apache Beam?

Beam is a programming model  
to build **batch and streaming** data pipelines



**Read** your data on-prem  
or in the cloud

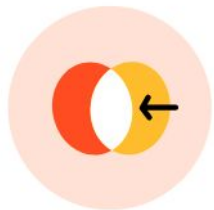


**Define** your business  
logic either for Batch or  
Streaming use cases



**Write** your derived data  
to the most popular  
destinations

# Why Apache Beam?



## Unified

Use a single programming model for both batch and streaming use cases.



## Extensible

Write and share new SDKs, IO connectors, and transformation libraries.



## Portable

Execute pipelines on multiple execution environments.



## Open Source

Community-based development and support to help evolve your application and use cases.

# Apache Beam: A Thriving Ecosystem

Spring 2019



Technology of the year  
2019 Awards

Summer 2020



Top 5 project in 4  
categories  
Significant Developer  
Interest  
4 years in a row!

—  
Top 5 Apache repositories by number of commits

—  
Top 5 most active mailing lists (user@ + dev@)

—  
Top 5 most active Apache sources --clones

—  
Top 5 most active Apache sources --visits

# Welcome to the first edition of Beam College!

A training program to learn about advanced data processing concepts and how to build **portable batch & streaming** pipelines using Apache Beam.

## 5 Days and 25h of content

**April 7, 2021:** The distributed data processing landscape, and Apache Beam under the hood

**April 8, 2021:** Advanced distributed data processing

**April 9, 2021:** Scaling and productionalizing your Beam pipelines

**April 15, 2021:** Strategies for performance and cost optimization

**April 16, 2021:** Best practices for debugging Beam pipelines

# Making the most out of this training

Tons of great content and valuable information **SO DON'T FORGET TO:**

- Try for yourself first-hand with the **code samples** - practice makes perfect
- **Ask questions** on slack - we're here to help!
- **Connect** with the Beam community

## Our vision for the future

- Always fresh and up to date content
- Add resources for other runners, new IO connectors, sinks and sources
- Build more use case based code samples and demos



# Housekeeping Announcements

Join the Slack Channel: <https://sg1.run/bcslack>

Clone the GitHub repo with slides and code samples: <https://sg1.run/bcgithub>

Watch the recorded sessions in our YouTube channel: <https://sg1.run/bcyoutube>

Help us make this training better by completing the survey: <https://sg1.run/bcsurvey>

**Invite your friends to join!**

# Coming up Next

Apache Beam in the data processing life cycle

Apache Beam use cases:

- Renault: Industry Data Management 4.0
- Alex Van Boxel: Give me my data now! Beam use-cases in review

# Welcome!

