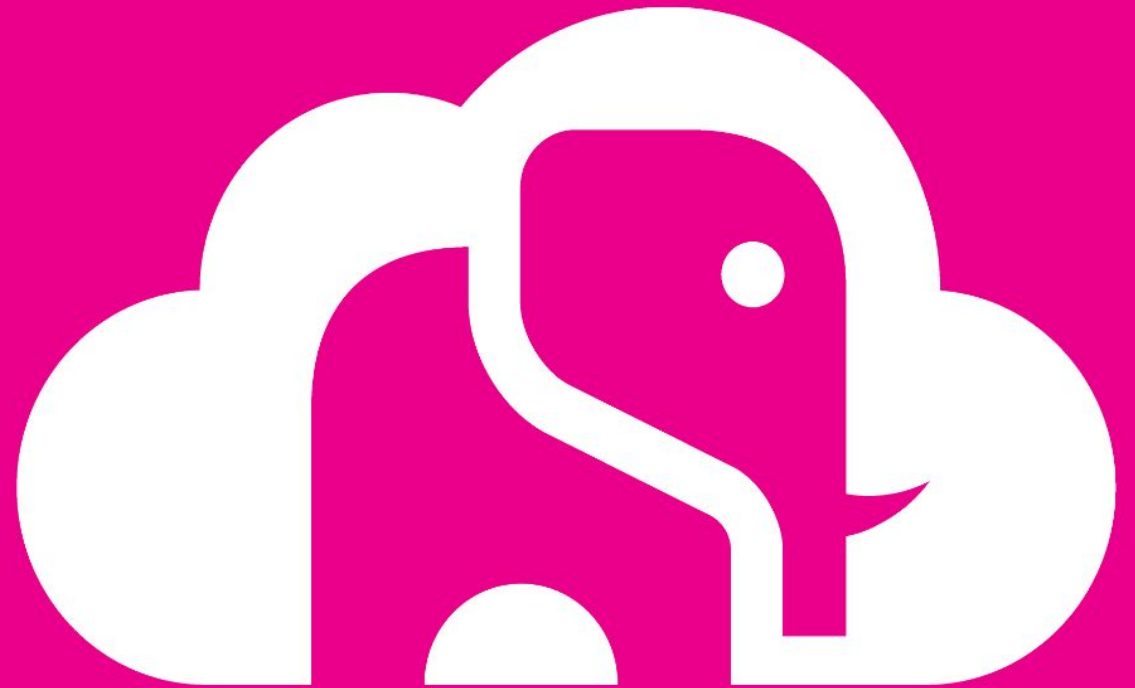# IF DATAOPS IS NOT DEVOPS PLUS DATA, WHAT IS IT, EXACTLY?

Doug Ortiz
Senior Postgres DevOps Engineer

June 2022

EDB

WELCOME

Housekeeping Items

Slides and recording will be
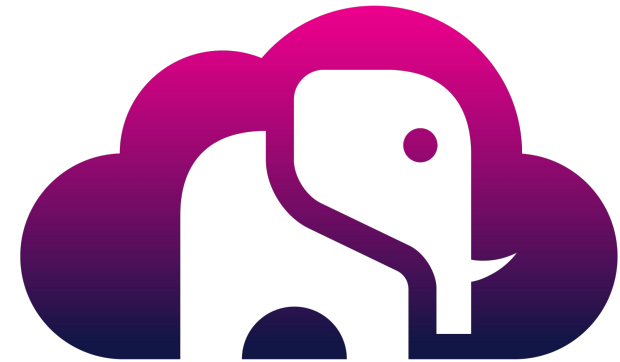available within 24 hours

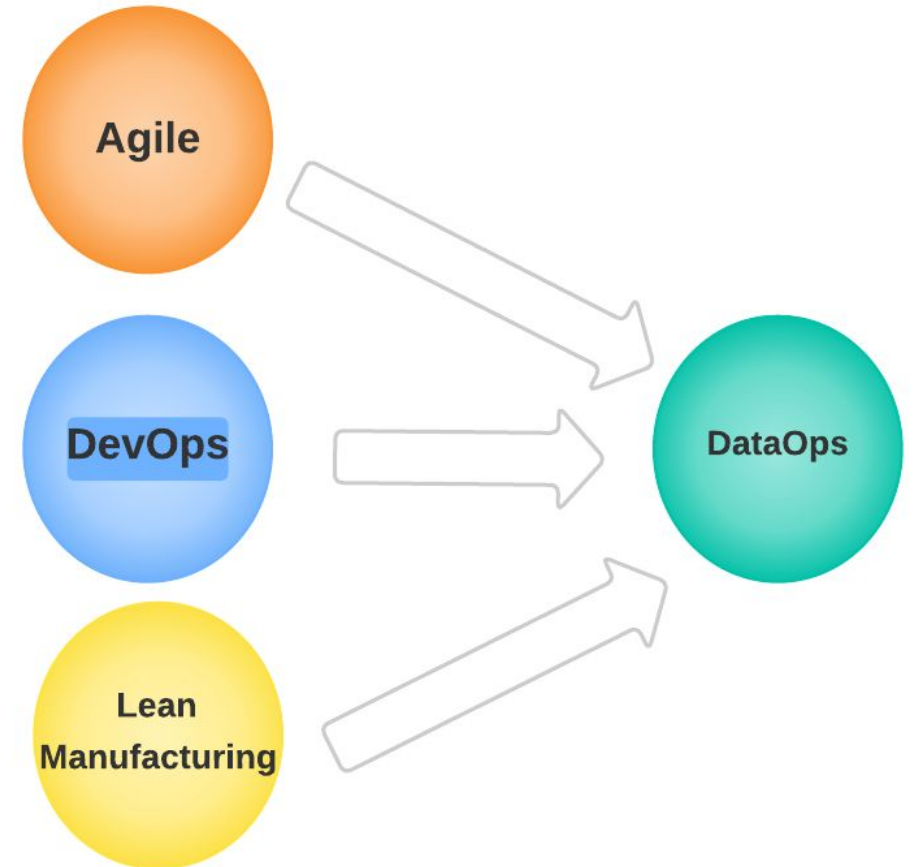**?**

Questions will be
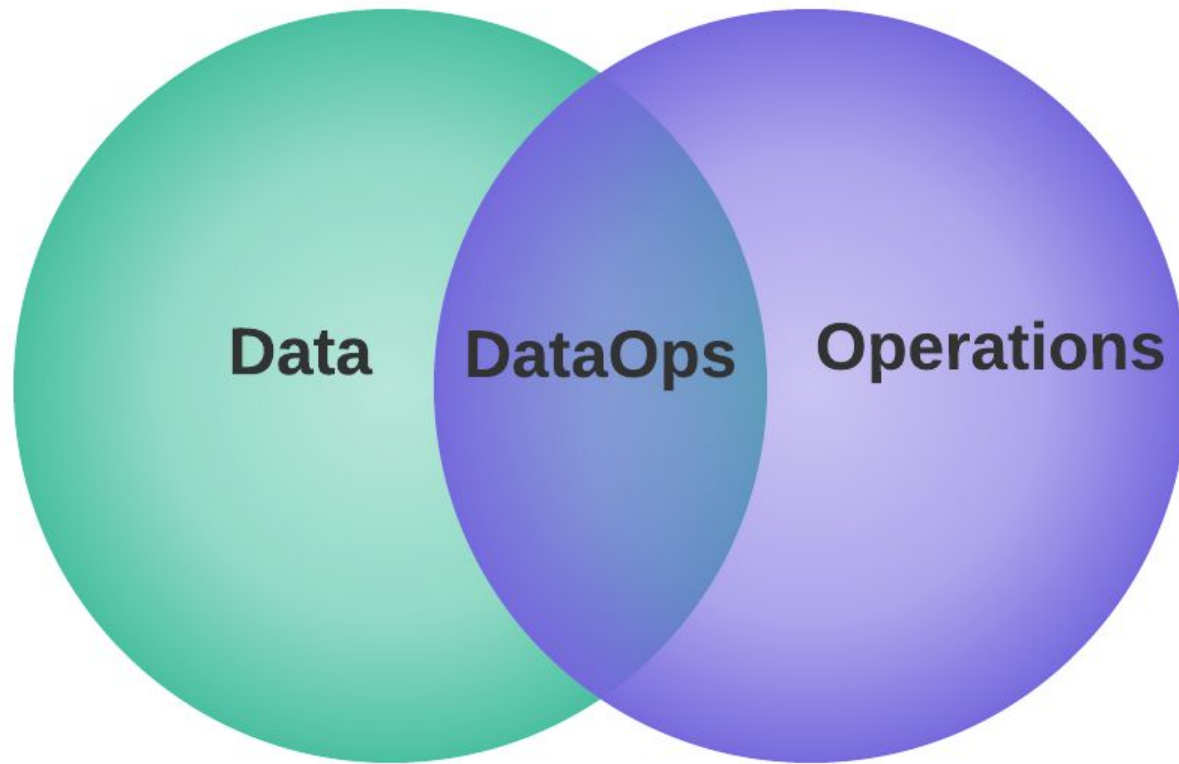answered at the end

AGENDA

If DataOps != DevOps + data, then what is it, exactly?

- What is DataOps
  - What is Lean manufacturing
  - Principles of Lean manufacturing
- DevOps
  - Core Values
  - Principles
- Why DataOps is not the same as != DevOps + Data
- Why Implement DataOps
- Benefits of utilizing DataOps
- Principles

DATAOPS - WHERE THE TERM COMES FROM

LEAN MANUFACTURING

What is Lean Manufacturing

- **Elimination of waste to continually improve a process**
- Derived from Toyota's 1930 operating model "The Toyota Way"
- Coined in 1988 John Krafcik
- Defined in 1996 by Professor James Womack and Daniel Jones

LEAN MANUFACTURING - PRINCIPLES

Defined Principles

- Precisely specify value by specific product
- Identify the value stream for each product
- Make value flow with interruptions
- Let customer pull value from the producer, and pursue perfection

DATAOPS

What is DataOps

- Is the combination of Agile development, DevOps, and Lean Manufacturing (*elimination of waste to continually improve a process*)

- Seeks to provide the tools, processes, and automation to handle increase in data demands

- Combination of data analytics development and data operations

- **Aims to increase velocity, reliability, and quality of data analytics**

- Set of practices, processes, and technologies that combine an integrated, and process-oriented perspective on data with automation and methods from agile software engineering to improve quality, speed and collaboration and promote a culture of continuous improvement in the area of data analytics

DEVOPS - CORE VALUES

What DevOps is all about. Which one is it? CALM, CALMS or CLAMS?

- Culture
- Automation
-
- Measurable
- Shareable

Lean

- Not originally part of the core values
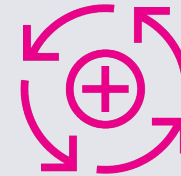- Added recently to expand and enhance the DevOps Core Values

DEVOPS - PRINCIPLES

Covered by 3 important aspects between Developers and Operations
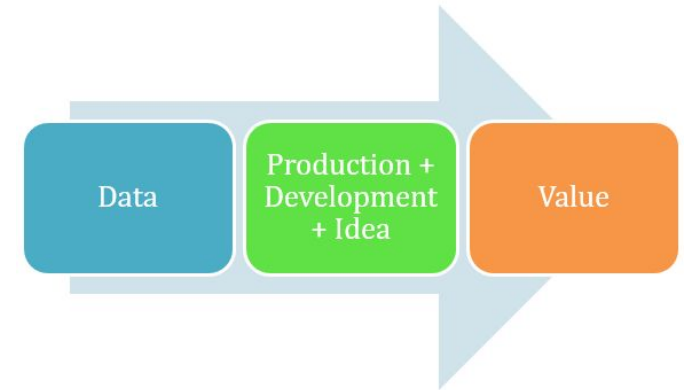
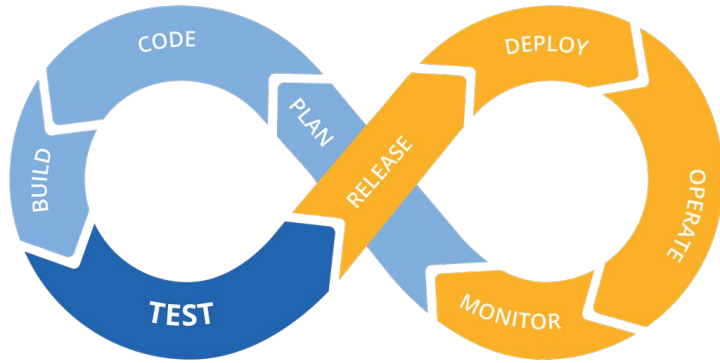**Systems thinking**

Amplify feedback loops

Continuous Experimentation

DATAOPS != DEVOPS + DATA

Flows

DATAOPS != DEVOPS + DATA

Resources & Tools

DevOps

- Software engineers

- Multiple

  - Coding languages

  - Tools

  - Pipelines

DataOps

- Data Scientists

- Engineers

- Analysts

- Tools

  - Pipelines

DATAOPS != DEVOPS + DATA

Processes

DevOps

- Develop

- Continuous Integration

  - Build

  - Test

- Continuous Deployment

  - Deploy

- Run

DataOps

- Sandbox Management

- Develop

- Continuous Integration

  - Orchestrate

  - Test

- Continuous Deployment
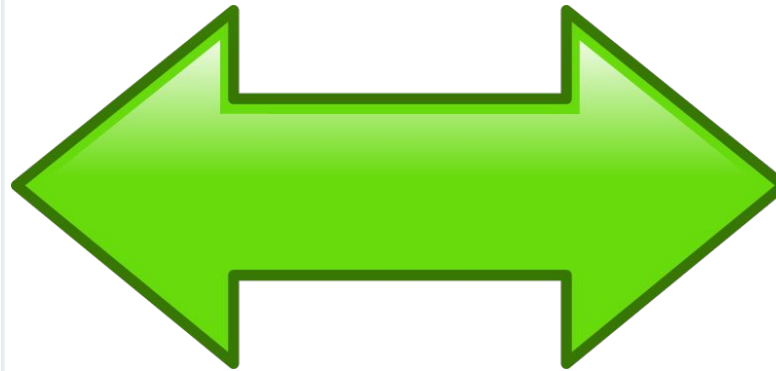
  - Deploy

- Orchestrate

- Monitor

DATAOPS - WHY IMPLEMENT

Connects the organization

Development

- Analysts
- Data Scientists
- Data Engineers
- Data Architects
- Developers

Cycles of innovation

- Centralized production teams
- Centralized Data Engineering, Analytics, Science, Governance development teams
- Groups self-service tools

Operations

- Production Team
- Monitoring
- Customers

DATAOPS - STEPS TO IMPLEMENT

Implementation Steps

- Add data and logic tests

- Use a version control system

- Branch and merge

- Use multiple environments

- Reuse & containerize

- Parameterize your processing

- Work without fear or heroism

DATAOPS - BENEFITS

- Reduce time to insight

- Improve data analytics quality

- Higher quality of data releases

- Lower the marginal cost to ask the next business question

- Promote team efficiency through agile process, reuse, and refactor

- Application of DevOps culture, and philosophies to Data

DATAOPS - PRINCIPLES

- Continually satisfy your customer

- Value working analytics

- Embrace change

- It's a team sport

- Daily interactions

- Self-Organize

- Reduce Heroism

- Reflect

- Analytics is Code

- Orchestrate

- Make it reproducible

- Disposable environments

- Simplicity

- Analytics is manufacturing

- Quality is paramount

- Monitor quality and performance

- Reuse

- Improve cycle times

DATAOPS - TOOLS

- Source Control for Scripts
  - Tables, Views, Indexes, Constraints, Stored Procedures, Functions, Triggers, and Database Configuration
- Schema Source Control
- Data Change Scripts - Data Manipulation
  - Database Objects
  - Data with Database Objects
- Version control data and/or a database
  - Liquibase
  - Flyway
- Data Testing
  - Faker

- Provisioning and configuration of Database Clusters
  - Terraform
  - Ansible
  - Puppet
  - EDB postgres-deployment
  - EDB edb-ansible

- DevOps Tools for Databases
  - Containerization - Docker, and PodMan
  - Orchestrators - Kubernetes, and Openshift
  - Source Control
  - Job Scheduling

Walk through of DataOps example

Takeaways

DATAOPS - TAKEAWAYS

Recommendations

- Embrace the DataOps Manifesto - www.dataopsmanifest.org/

- Recognize the differences between DevOps and DataOps

  - Flows

  - Resources

  - Tools

  - Processes

- Architect/Design an implementation that fits your organizations needs

- Assess which tools can be leveraged for your DataOps implementation

- Welcome errors and failures as part of the process

- Revisit the DevOps, and DataOps principles, tailor them to your liking
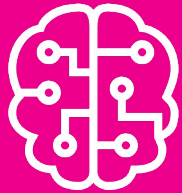
DATAOPS

Useful Links

- Embrace the DataOps Manifesto - www.dataopsmanifest.org/

- GitHub and Demo Code

    - https://github.com/EnterpriseDB/blogpostings/tree/main/DataOps/dataops-demo

Q&A

THANK YOU

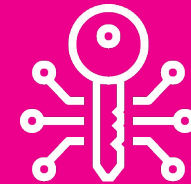## BigAnimal: Faster, safer, smarter, better

Postgres Expertise

Expertise beyond the generalist cloud provider; we help steer the database roadmap and patch its bugs

Oracle Compatibility

Leave Oracle and further your cloud journey with a fully managed Postgres service

Greater transparency and control

BigAnimal runs in your Azure account and leverages your existing discounts

Curious? Request a free trial today!
https://resources.biganimal.com/cloud-postgresql-trial