Apache Airflow at Wise.
Money without borders.
Wise today

10 million customers around the world

£4.5 billion processed for customers every month

£1 billion saved for our customers every year

2,200 staff from 70+ nationalities
Alexandra Abbas
Machine Learning Engineer
@alexandraabbas
What do we use Airflow for?

- **Treasury team**
  Monitor liquidity and cash buffer in our system

- **Finance team**
  Generate Google Sheets reports to BI stakeholders

- **Fraud team**
  Fraud detection for different payment methods
Our data stack

- **Kubernetes**
  Custom HTTP service for serving predictions

- **Amazon Sagemaker**
  Data science experiments and model training

- **Amazon EMR**
  Data processing and model training

- **Snowflake**
  Data warehouse and SQL data analysis
Requirements for Machine Learning Platform

- **Security**
  Minimise vulnerabilities and the chance of breach

- **Segregation of data**
  GDPR, partner requirements, restricted access

- **Autonomy for teams**
  Teams manage their own code base
Our Machine Learning Platform

- Prediction service
- Airflow
- Sagemaker

Marketing team

Support team

Secure AWS environment
Evolution of Airflow at Wise

- Single repo
  - Single instance
- Single repo
  - Multiple instances
- Multiple repos
  - Multiple instances

Time
Our Airflow architecture

- Proxy
- Master EC2
  - Web server
  - Scheduler
  - Flower
  - Redis
- Worker EC2(s)
  - Worker
  - Exporter
- Artifactory
- Amazon RDS
Airflow infrastructure management

- Amazon RDS
- EC2
- S3 buckets
- IAM role
- Secrets Manager
- Terraform
Airflow infrastructure management

- Install Docker
- Setup airflow user
- Create cronjobs
- Pull secrets
- Start containers

EC2

Userdata
<table>
<thead>
<tr>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Easy to create a new instance</td>
<td>Scalability</td>
</tr>
<tr>
<td>Segregated environments</td>
<td>High availability</td>
</tr>
<tr>
<td>Secure environment</td>
<td>No SLAs</td>
</tr>
</tbody>
</table>
Questions.