Big Data Analytics & ML

About us

10+
years serving

20+
million end customers

100+
projects

Our **Core Services**

UX Design

Product Development

Data Analytics & Machine Learning

Intelligent Automation

Our Team

1. Data Analysts

- Domain knowledge of multiple domains
- Expertise defining business use cases & KPIs with stakeholders
- Hands on experience in Descriptive, predictive and prescriptive analytics
- Identify data sources, Attribute relationships mapping to KPIs

2. Data Engineers

- Experience in recommending optimal, low-cost tech stack for ETL
- Data lake for structured and unstructured data
- Skills in Spark, Airflow, SQL, Databricks, Snowflake, Social media APIs
- Third-party SaaS product integration knowledge
- ETL on AWS, Azure, GCP and on-premise

3. Visualization Experts

- Decide on apt dashboard visualization for quick decision-making
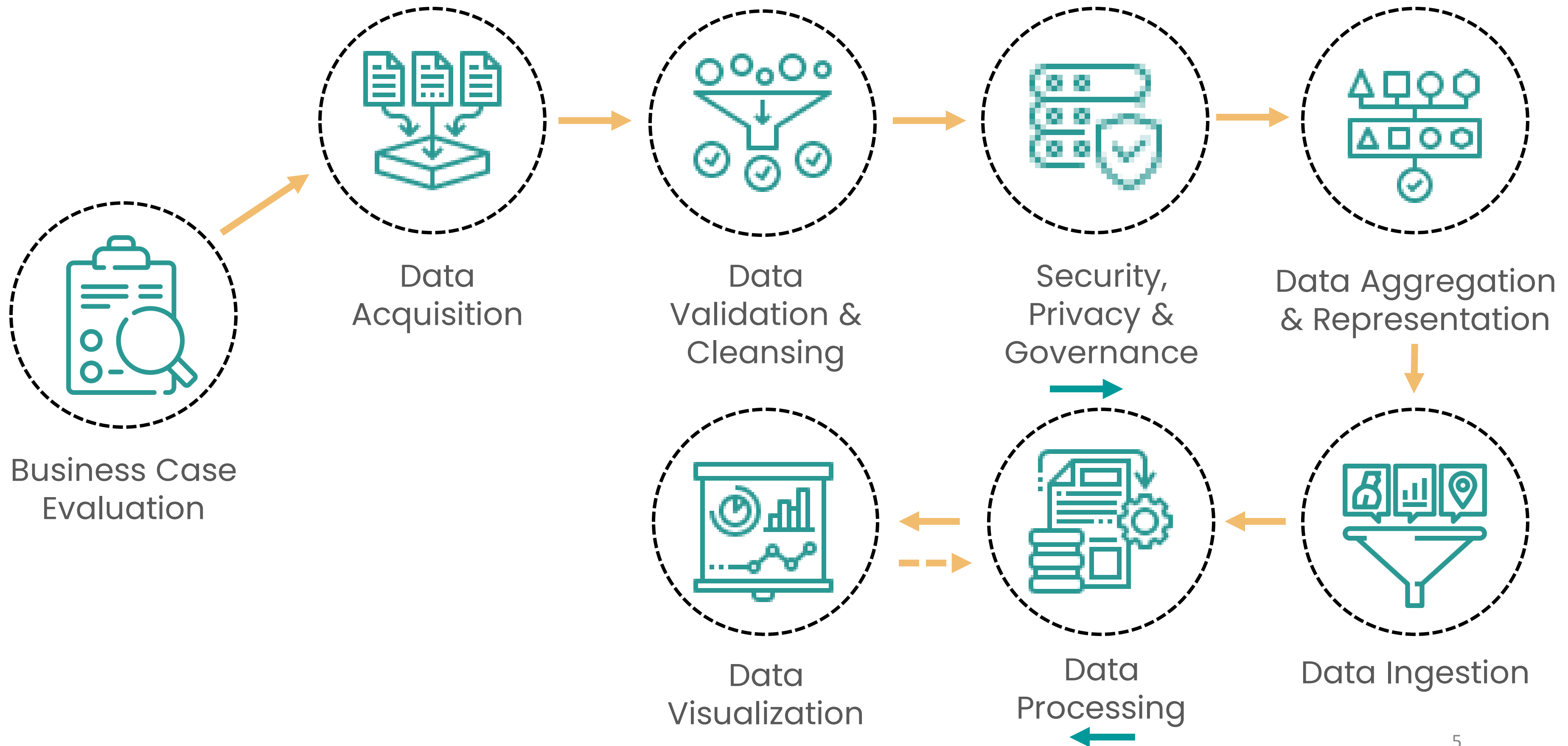- PowerBI, Tableau, Redash, Plotly, Quicksight, Superset

Our Team

4. Data Scientists

- Data normalization and cleaning
- Feature Engineering
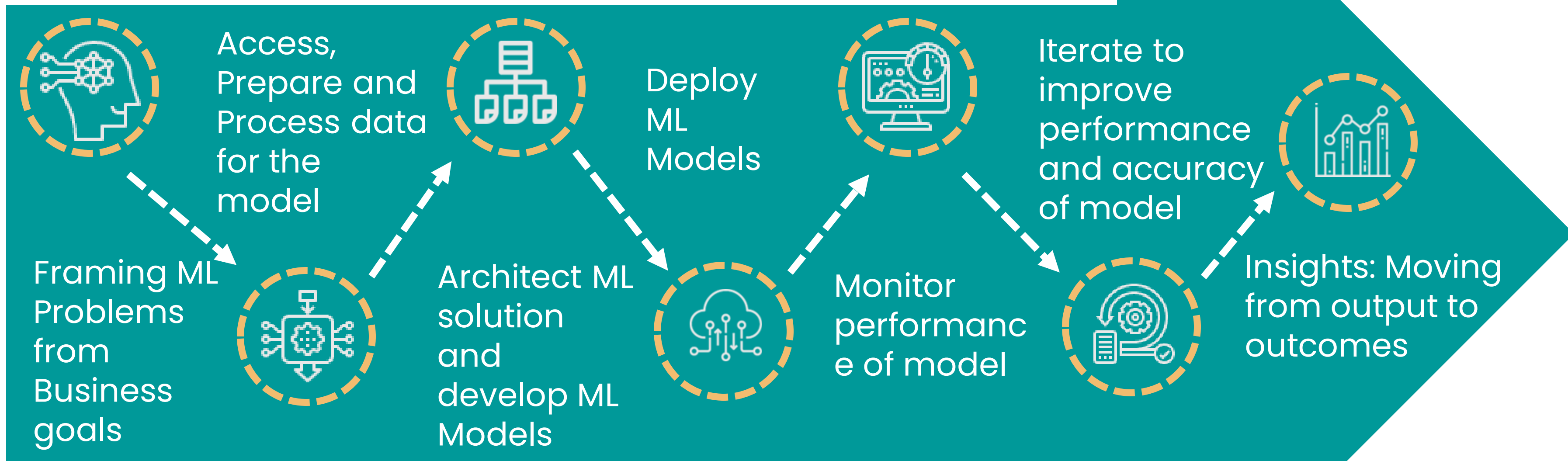- Unsupervised and Supervised Learning Algorithm Selection

5. Machine Learning Engineers

- Deep Learning Algorithms implementation using TensorFlow, Yolo, BERT Frameworks
- Unstructured data interpretation – Computer Vision, NLP
- MLOps pipeline for continuous improving of ML model accuracy
- High performance deployments over GPUs
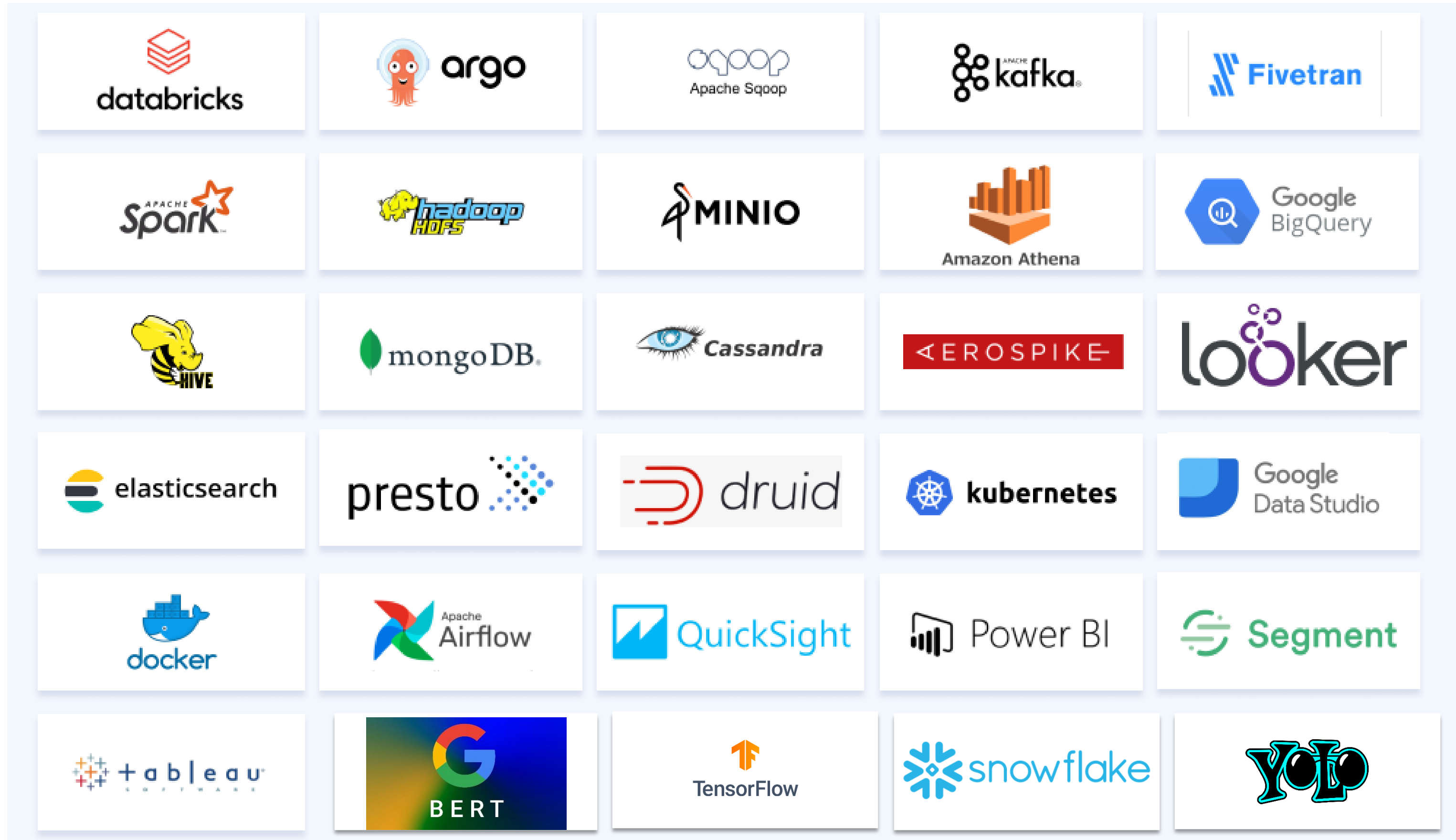- Knowledge of Cloud ML Services like Sagemaker, AIBuilder etc.

Data Analytics Process

Business Case Evaluation → Data Acquisition → Data Validation & Cleansing → Security, Privacy & Governance → Data Aggregation & Representation → Data Ingestion → Data Processing → Data Visualization

ML Lifecycle

Access, Prepare and Process data for the model

Deploy ML Models

Iterate to improve performance and accuracy of model

Framing ML Problems from Business goals

Architect ML solution and develop ML Models

Monitor performance of model

Insights: Moving from output to outcomes

Technologies

Domains we Serve

Finance

Government

Healthcare

Agriculture

Logistics & Transportation

IoT Sensor Data

Sales & Marketing

Gaming

Our Pillars

People

Technology

Process

Domain Knowledge
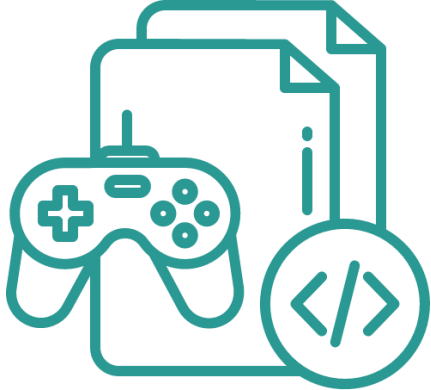
Our Analytics & ML Services
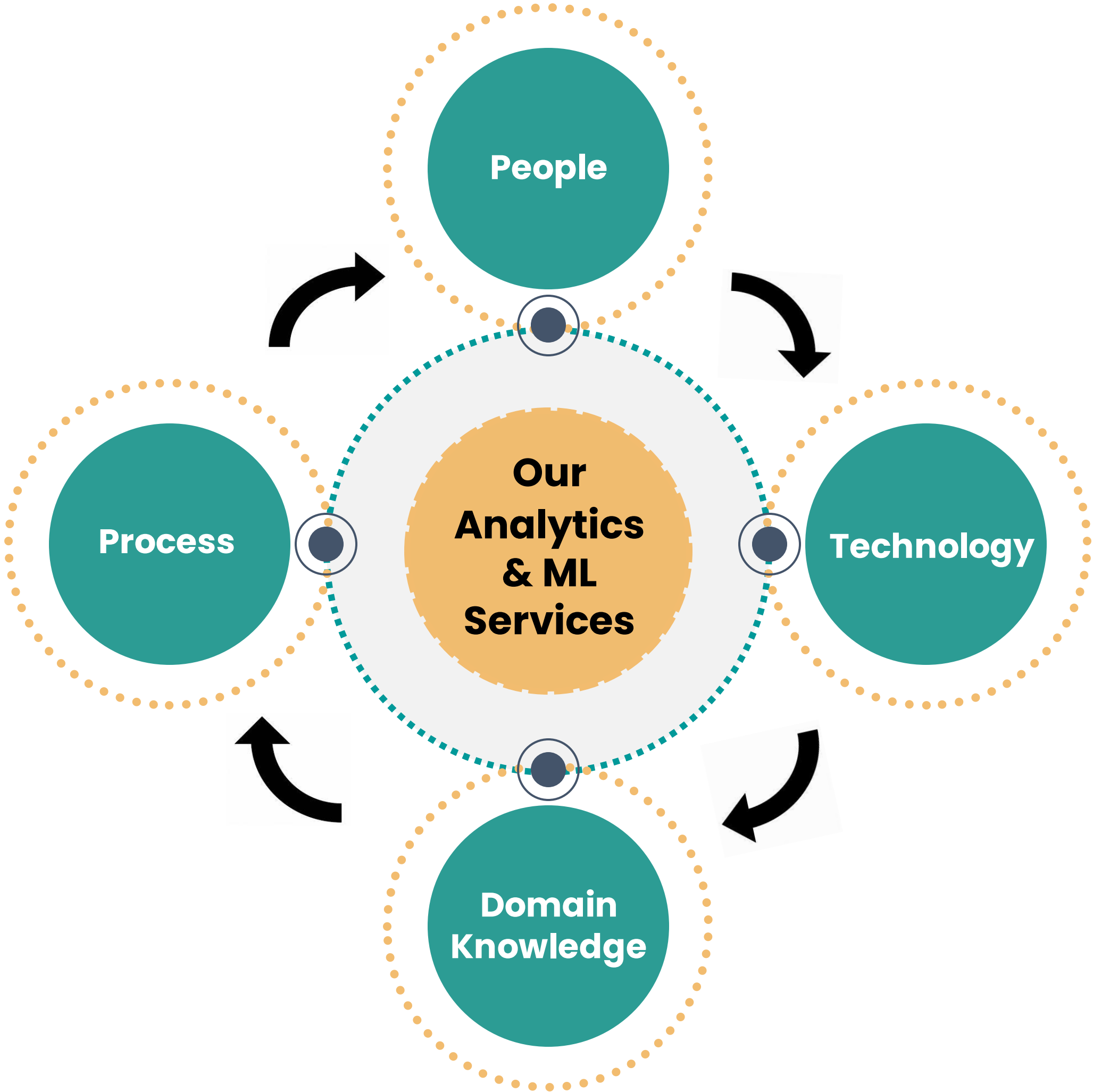
Journey to ROI

Understanding Requirements

- Articulating business needs and estimated impact
- Creating a base with a vision of how the solution will be beneficial for business

Data Acquisition

- Gathering Data from internal and external sources
- Combine technical and business expertise to unlock the value of big data
- Enhance data by filtering out unproductive information

Data Engineering

Deriving analytics through the right combination of quantitative techniques and cutting edge technologies

Modelling Insights

Generating actionable insights to improve customer acquisition, retention, and efficiency

Workflow Integration

- Integrating insights into decision making to automate processes
- Deploying latest tools for converting insights into actions to drive outcomes

Outcomes

- Operational excellence
- Increased security
- Enhanced customer retention
- Expanding market share

USE CASES

Sales & Marketing Funnel

- Ingest customer data from various sources: CRM, Call Centre and social media platforms
- A single source-of-truth for 360 degree view of customer journey
- Extract insights and develop reports for the sales and marketing departments
- Visualize insights on dashboards for actionable decisions
- Predict future sales, conversions
- Allocate optimal ad-spend budget to various marketing channels for better ROI and conversions
- Tech - Fivetran, Astronomer, Databricks Deltalake, PySpark, Redash
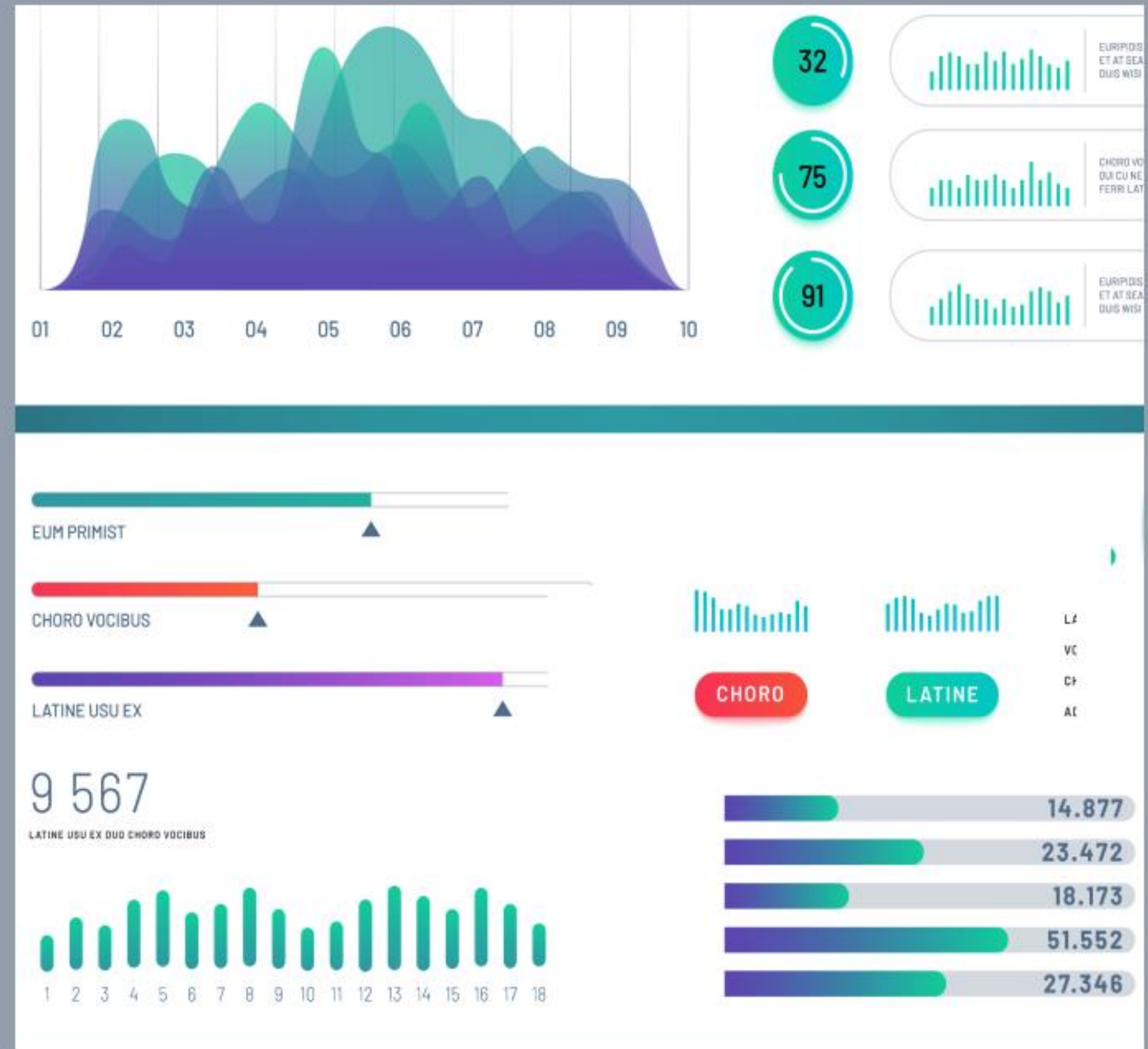
Rideshare Analytics

- Capture ride data from mobile apps of passengers and drivers
- Extract insights and develop reports for ride fulfilment, cancellations, demand, supply, locations, driver effectiveness, reviews, route, and traffic
- Visualize insights on dashboards for actionable decisions
- Predict passenger interests, optimal routes, and better prices
- Run targeted push and In-App ad campaigns with gamification for discounts, referrals, promotions based on analytics
- ETL pipeline with Kafka, Postgres, D3JS
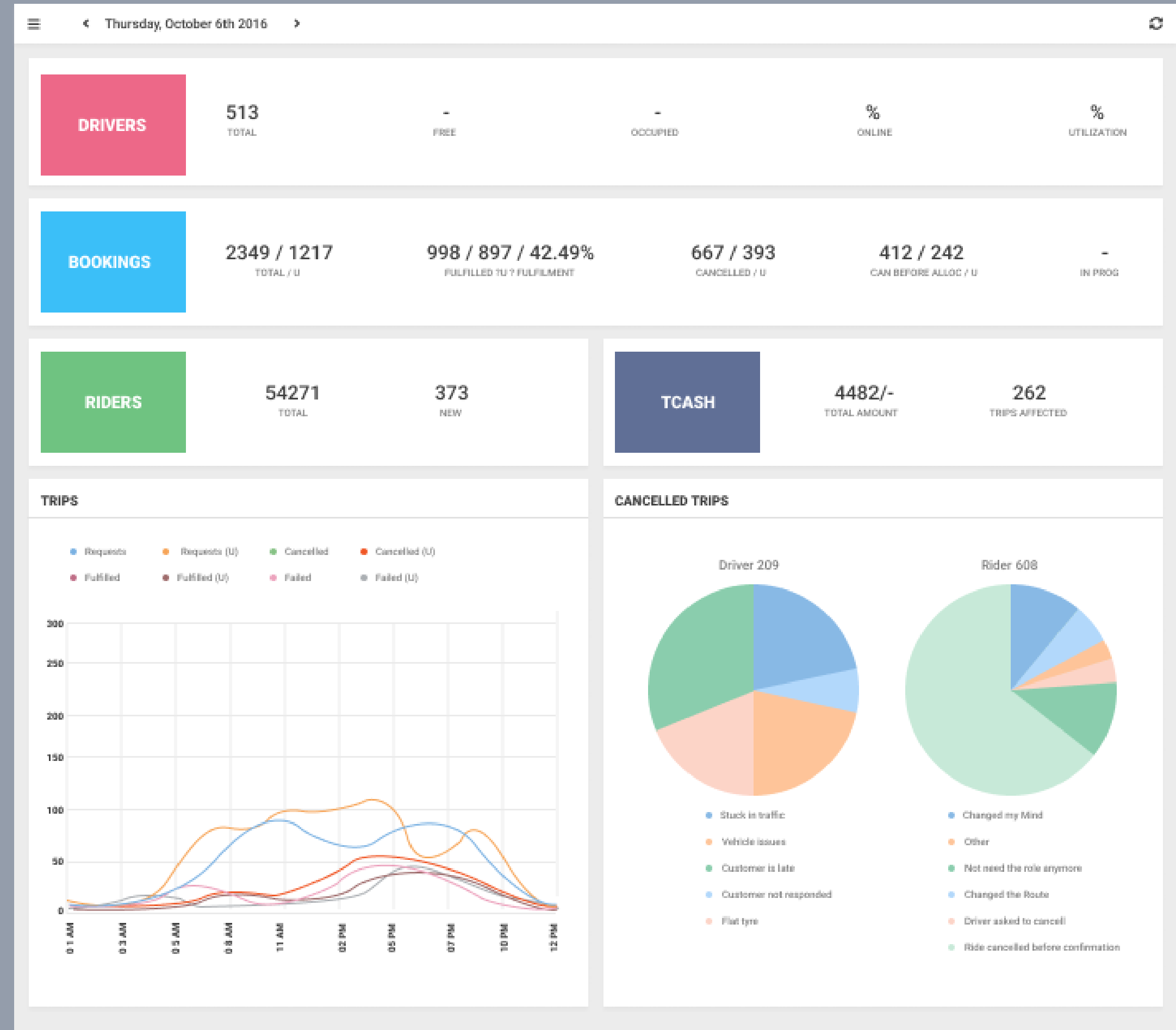
Document understanding

- NLP for named entities, context recognition for human-language based interpretation
- Domain specific Taxonomy and Ontology for human-knowledge based interpretation
- Combination of vision, language, knowledge and rules for human-like reproduction of document insights
- Document Classification
- Field Extraction: Checkbox, Layout – Title, List, Text, Image, Table -Row, Column, Cell, Signature, Date, Stamp, Barcode
- Predictive analytics on corrective operations data
- Celery, Kuberenetes, RabbitMQ, Redis, Postgres

Healthcare

- Ingest unstructured and structured data from Electronic Health Record (EHR), Claims and Enrollment
- ETL to cleanup, process data
- Sqoop, S3 with Parquet storage, AWS Athena OLAP cube, PySpark for processing, Quicksight for analytics visualization, SageMaker for model building, ML with Tensorflow
- Patient lifetime visit analysis, Disease prediction, Drug-effect analytics
- Patient population analytics, Infectious disease location-based trends, demographics/race/gender based analytics
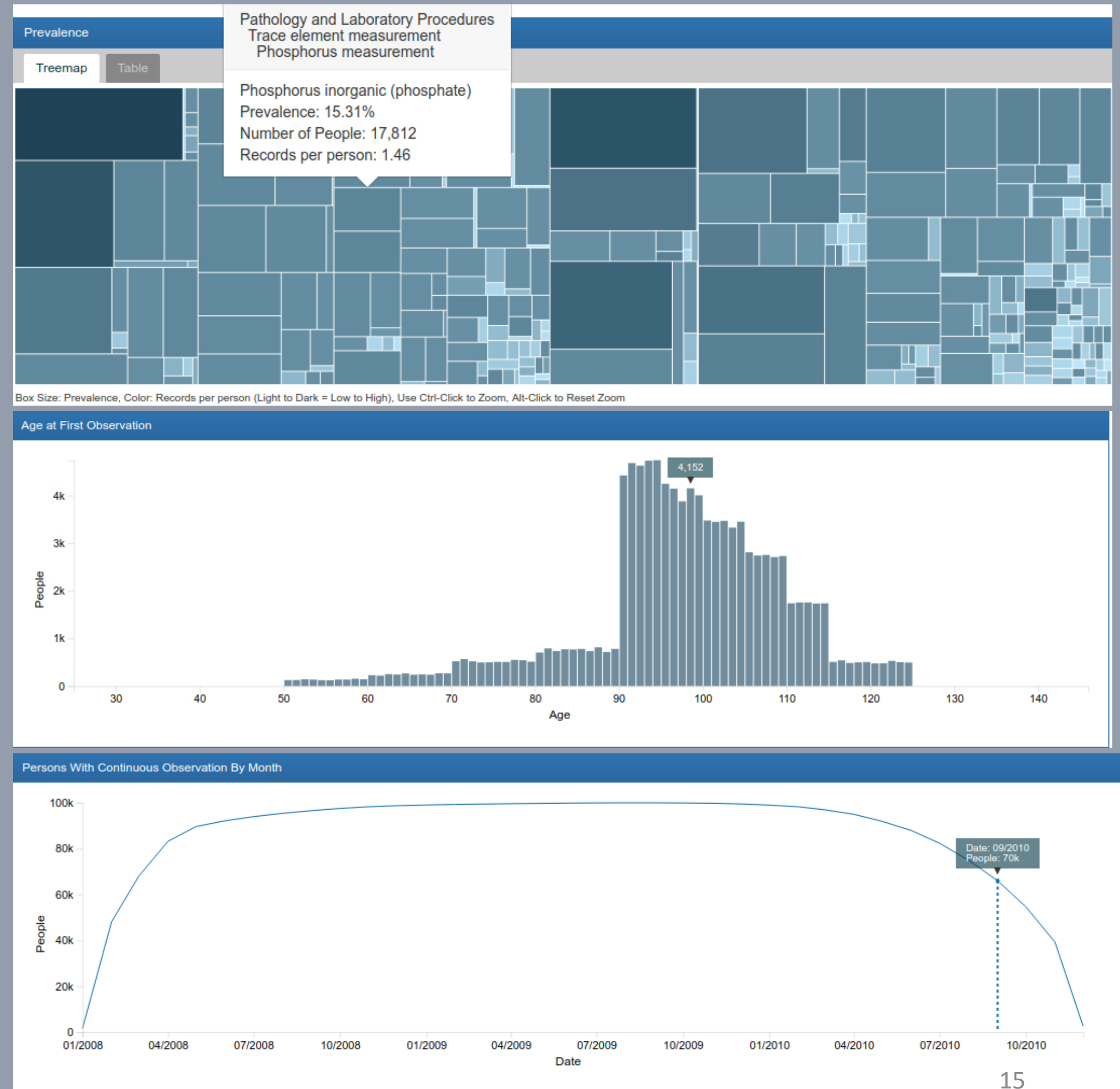
Gaming Analytics

- Capture Mobile Game Player and acquisition channel data – profile, game and click events, location, device details, social media
- Process to support real-time and batch analytics to improve acquisition, engagement, retention and monetization
- In-app and push Ad campaigns based on analytics
- Campaign effectiveness analytics and A/B testing analytics on player segments
- Game popularity, player behavior, ad effectiveness prediction, personalized game experience
- Real-time event processing pipeline using Kafka, Apache Flink, Postgres, D3JS
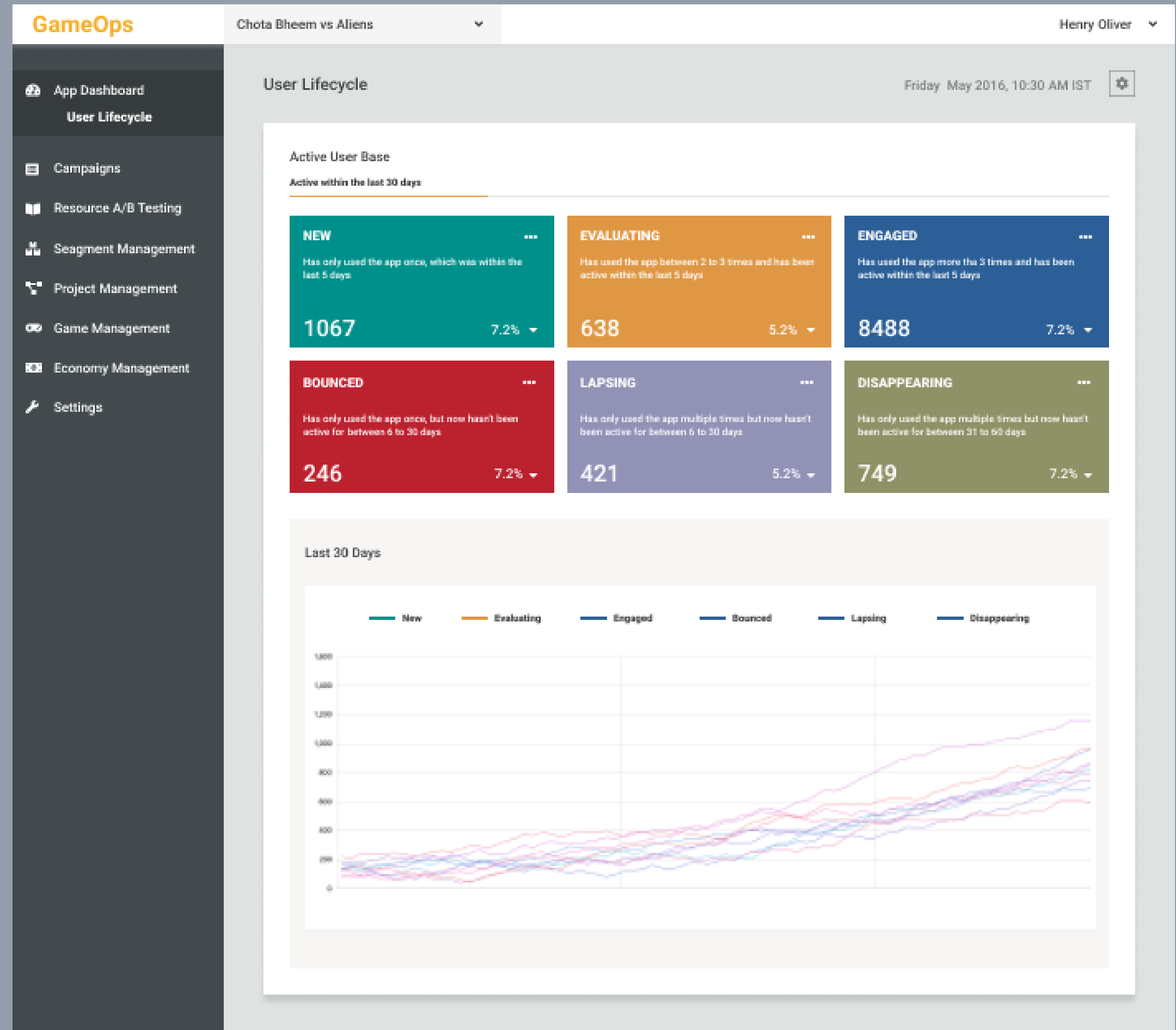
ISP Billing Analytics

- Capture data from subscribers, equipment, tickets, revenue, application events
- Process to support real-time and batch analytics
- Equipment depreciation analytics, Revenue Analytics, Subscriber retention analytics
- Predictive models of Machine learning for revenue prediction, capacity utilization and planning
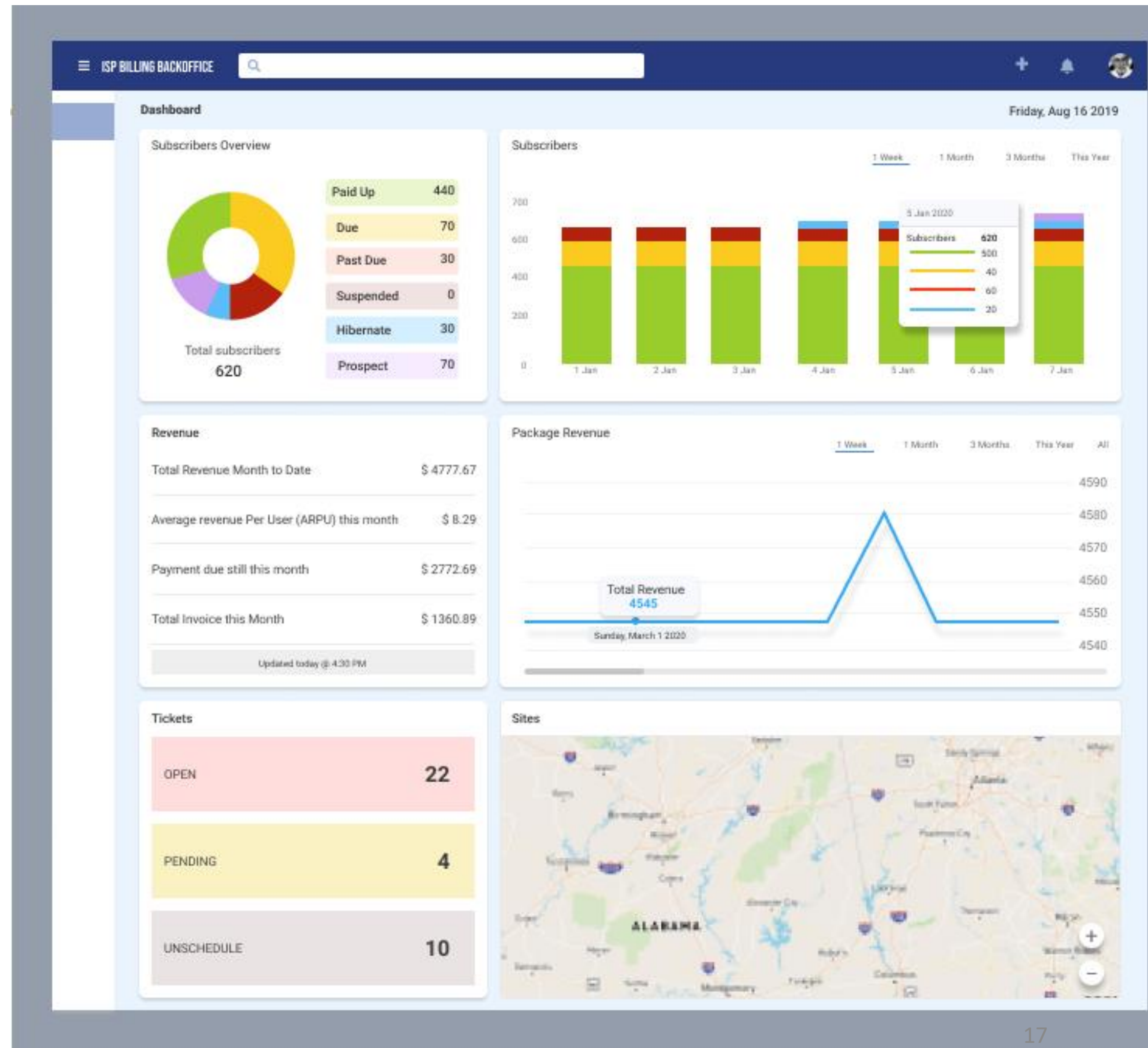
Mining Sensor Data Analytics

- Ingest real-time data from sensors embedded in mining equipment spread across multiple sites
- Sensor data includes temperature, speed, thickness, rotations etc.
- Track components that breached sensor threshold limits
- Drill-down visualization from customer to site to machine to component
- Breakdown prediction due to component failures for preventative maintenance
- Azure Functions, Azure PowerApps, SQL Server, D3js

Intelligent People & Asset Tracking

- Has fine-tuned location detection
- Hazardous and restricted area notification based on roles
- Attendance tracking with reports
- Locating all the people in a designated area during an emergency
- Real-time allocation over multiple sites for thousands of employees
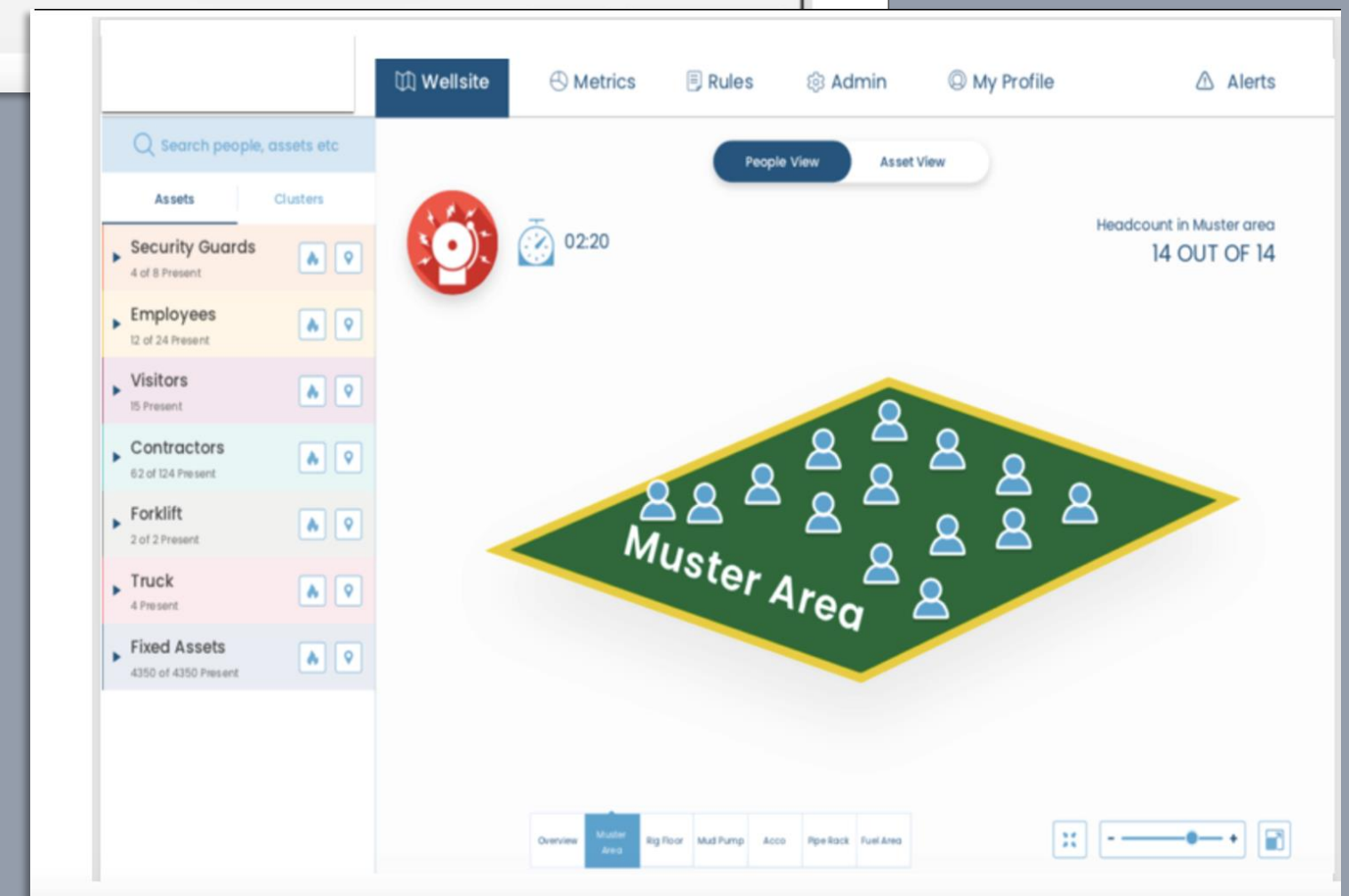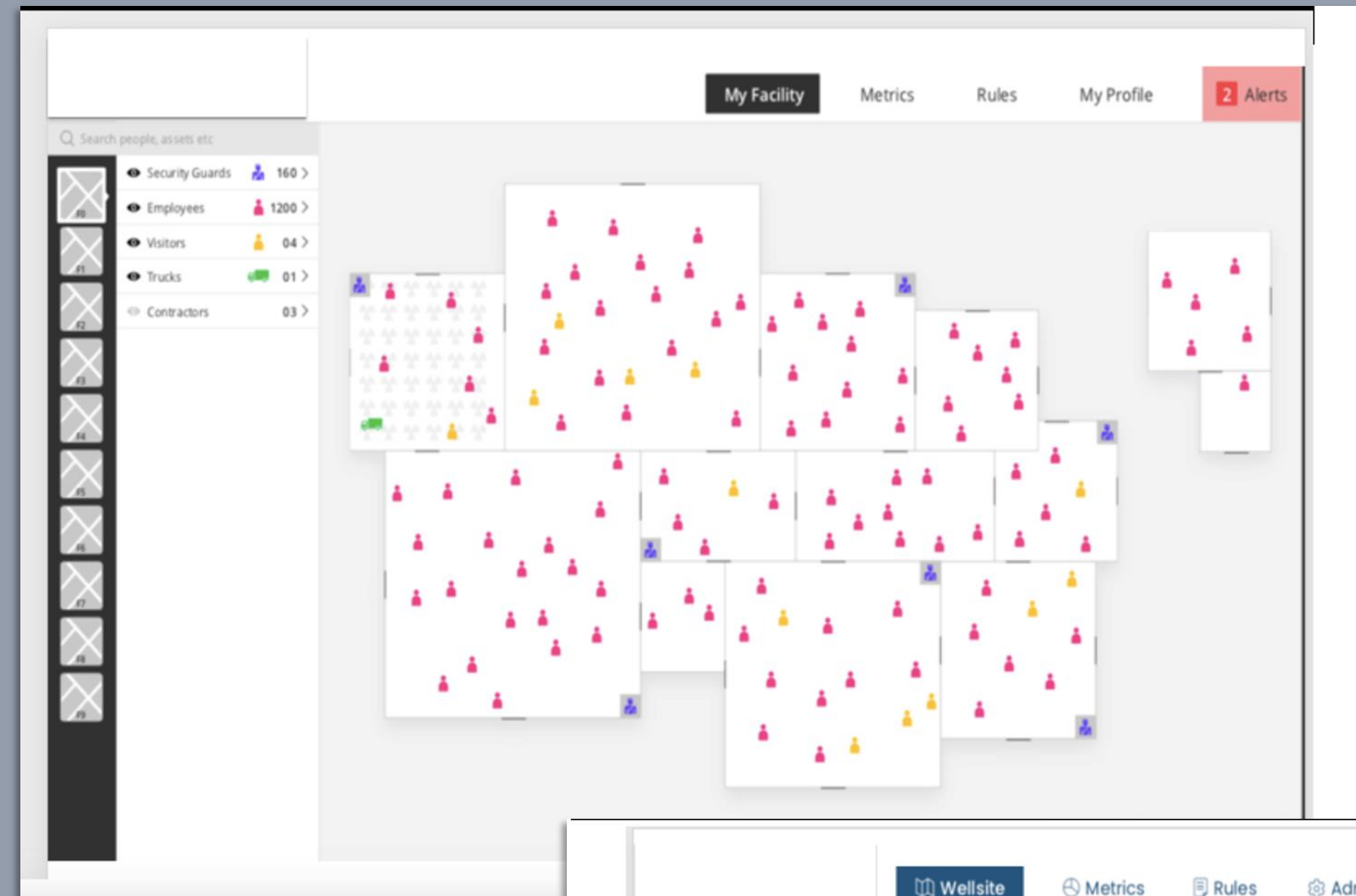
Precision Hydroponics

- Data collected from zones within indoor, climate-controlled hydroponic farm using temperature, humidity, light and nutrient IoT sensors
- Real-time sensor analytics and alerts
- Visualization and alerts on real-time dashboard
- Machine learning based prediction of optimum recipes for effective growth of multiple varieties of plants based on data collected from plant growth and sensors
- Automatic feedback loop to controllers based on predicted recipes for optimal plant growth
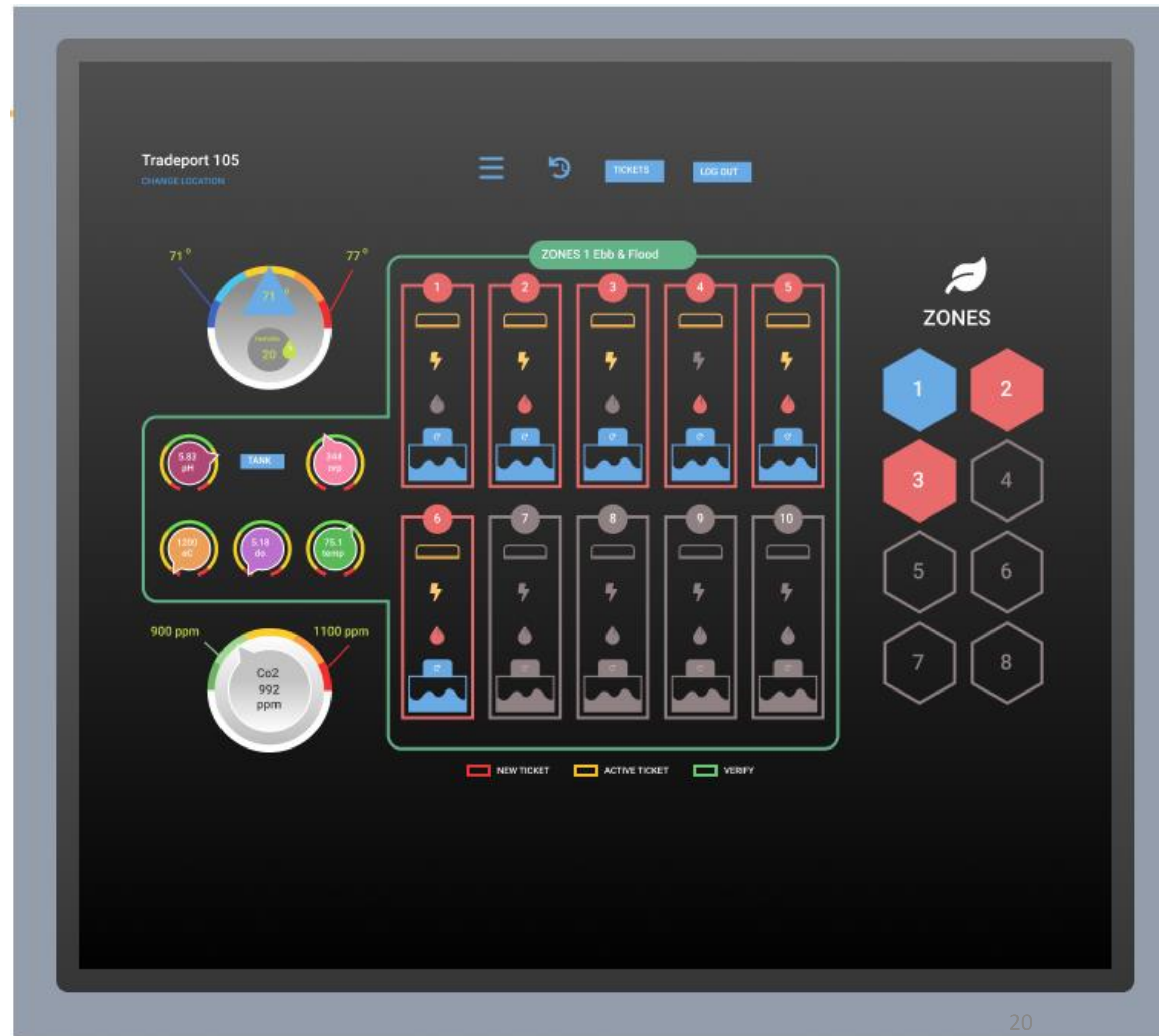
Agriculture Analytics

- Data collected from ground team surveys, government databases, farmer call centre logs, real-time market commodity prices, mobile apps, web portal and satellite photos
- Detect crop sowing patterns, crop diseases, regional crop and input sales
- Targeted SMS and Push campaigns for fertilizers, pesticides, seed recommendations and produce buy-back
- Supervised learning over satellite imagery to perform crop yield prediction using NDVI values
- Recommendations to streamline and optimize input and output extensions
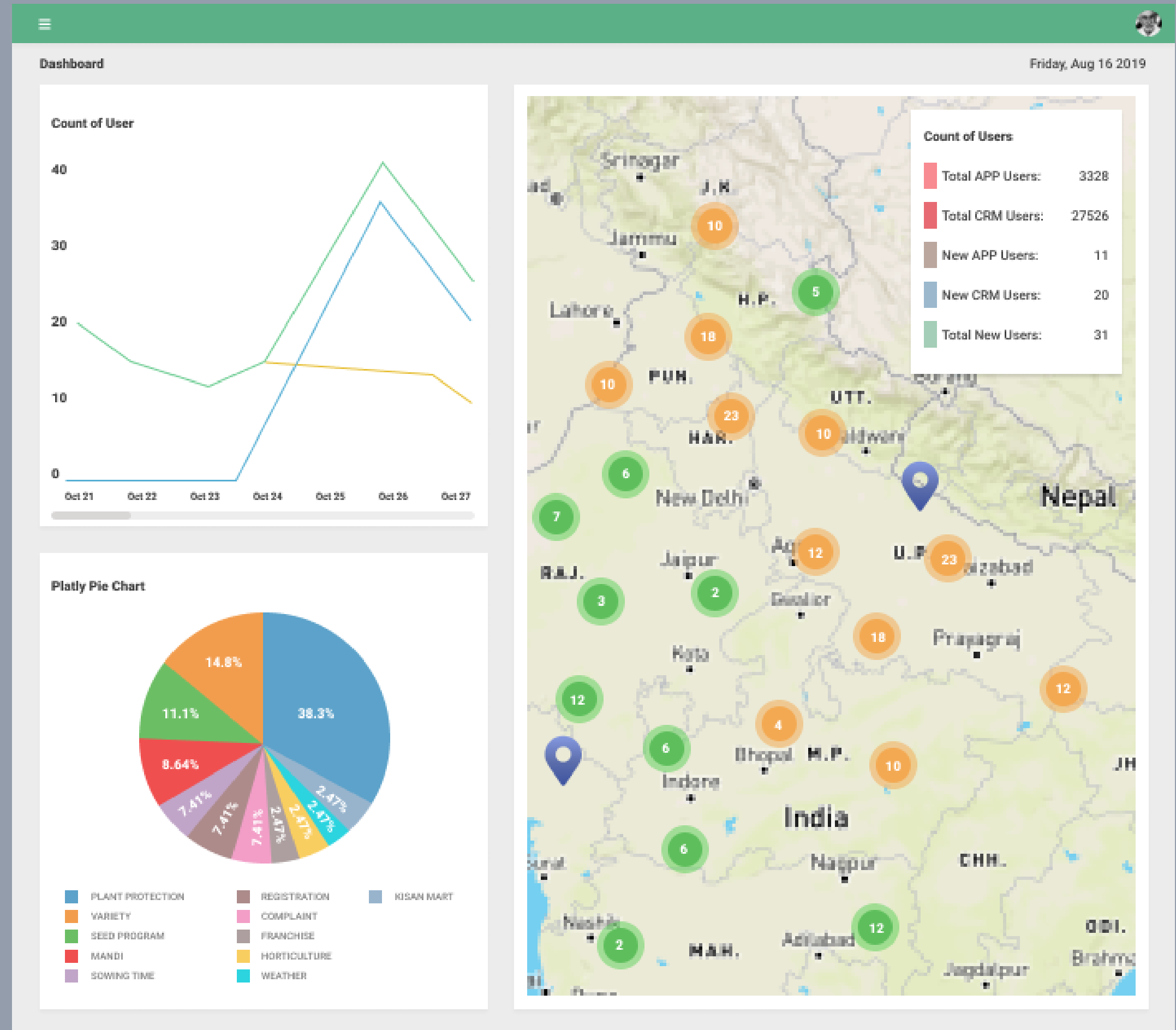
Govt. Service Analytics

- Collect data from databases of various social programs of the middle-east government: marriage fund, social security fund, office attendance, IT operations, Network operations

- Plot of attendance, geolocations, vacations, marriage fund applications, refunds, overlapping payments etc.

- Self-service BI Analytics using Superset.

- Prediction on budget, attendance, clustering of marriage fund applicants based on feature extraction, fraud detection using anomaly tracing
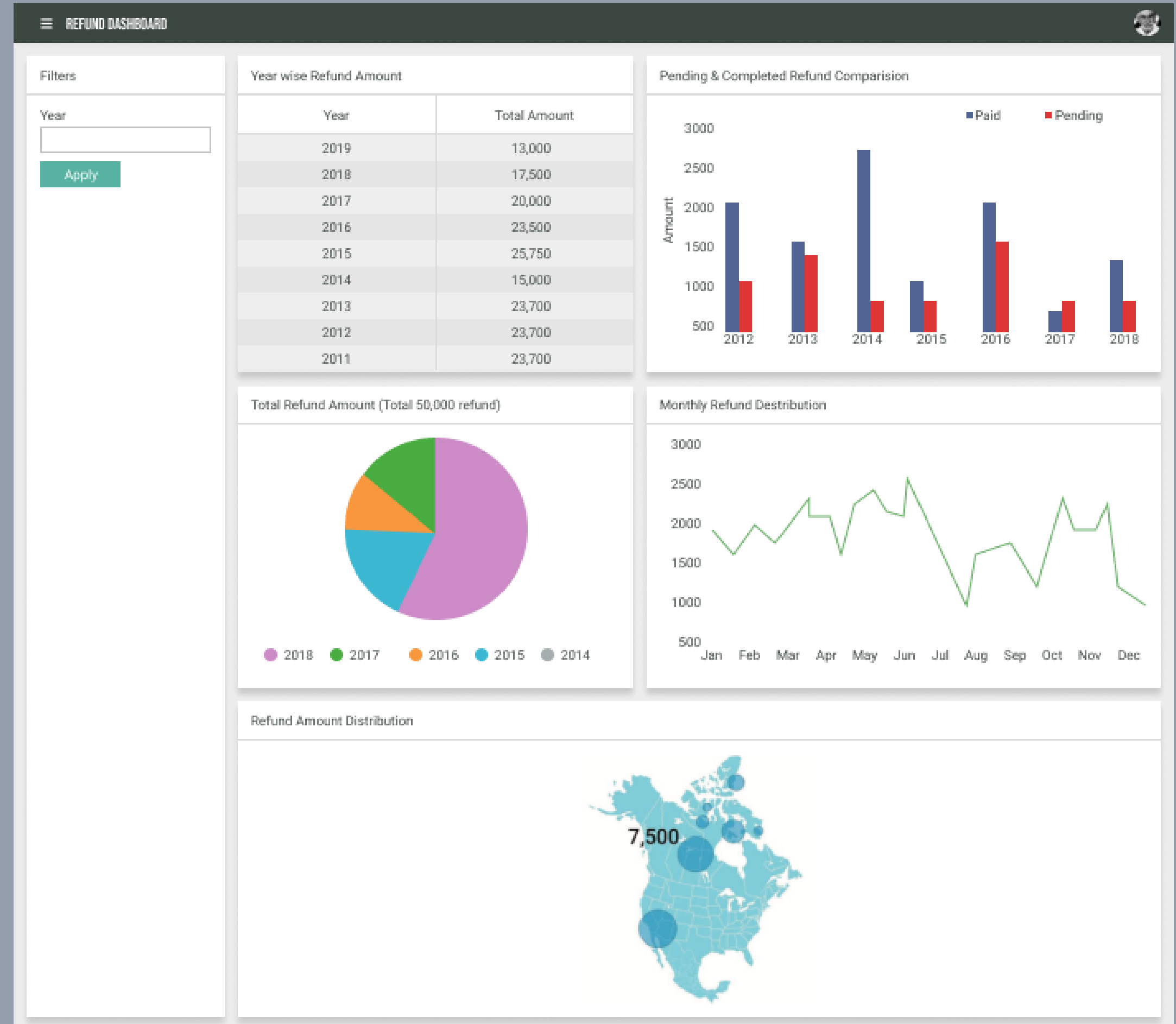
Why us?

- Proven expertise in **development of big data analytics** in various domains

- Build **lowest cost, scalable solution** using cloud or on-prem services

- Leverage pre-built components and products for accelerated development

- Transparent processes to drive overall **success**

- **Data Science and machine learning** capabilities for predictive and prescriptive analytics

- A **win-win partnership** based on trust

THANK YOU!

Hello World!

hello@47billion.com

47Billion