Using data analytics for rail – key challenges and ways to address them

Dr. R. Gopalakrishnan
General Manager
Centre for Railway Information Systems
Indian Railways
Outline

Focus Areas
Analytics life-cycle
General Challenges
Scaling up
Focus areas of business

- Revenue and customer satisfaction
- Efficiency and safety
- Investment planning
IR experience in analytics and data

• Passenger – seat capacity optimization, crew optimization
• Freight – travel time, demand forecasting, route optimization
• Network planning – level crossing and mid-section detentions
• Inventory management – materials master de-duplication
• Track maintenance – track structure
• Master data management
• Data lake
Analytics use case life-cycle

1. Finding the right Business Question
2. Locating Data Sources
3. Model Development
4. IT Integration
5. Consumption & change Management

Business user + consultant

Data scientist + IT team

Business user + IT team

Business Users

Data scientist and Engineers

Business Users
Single use-case problem statement (supply side)

1. Finding the right Business Question
2. Locating Data Sources
3. Model Development
4. IT Integration
5. Consumption & change Management

Undocumented area of business
Framing the right business question
Data scientist doesn’t know business processes

Collaboration - Business and data science team
Single use-case model development (supply side)

Exploring data sources
Interpreting data, data cleaning
Data Quality
Data access and sharing
First time problems - research oriented

Data Glossary/ Data catalog
Data augmentation – IOT/ sensors?
Tools and platforms as and when required
Single use-case consumption (demand side)

Explaining analytics to users
Data or analytics literacy
Business to lead implementation
Process changes sought by business

- UI/UX – ease of use
- Need based training
- business process documents
- user-manuals
Scaling up
### People
- Training
- New roles in organization

### Process
- Organization structure
- Steering committees
- Recruitment and training
- Procurement

### Technology
- Datalake
- Master Data management
- Dashboards
- Knowledge management portal
- Data management portal
Strategy, collaboration and training

Analytics strategy
Use case framework
Ideation workshops

Data Strategy
Cloud adoption
Rail appreciation course – for non-railway partners

Analytics, data literacy for end-users
Self-learning videos
Incentives for collaboration
AI scaling up cheat sheet

Multiple use-cases interwoven with each other
- Use case pipelines
- Concept of analytics architecture
- Analytics strategy

Data exploration, augmentation, quality
- Data Governance
- Data strategy
- Data glossary, automation

Data Analytics modeling
- Multiple delivery channels
- Domain for data scientists

IT integration
- Enterprise data lake, and ML ops
- Cloud adoption

Consumption
- Analytics, data literacy for end-users
- Self-learning videos
- Incentives for collaboration
Data Analytics Framework for Passenger Business

- **Waitlist prediction**
- **Passenger Demand Forecasts**
  - Dynamic Pricing
  - Ideal train profile
    - Standard Rake profiles
- **Passenger Train Time Table**
  - Optimal Rake Link
  - Optimal Loco Link
  - Optimal Crew Link

**Status: Completed | In progress | Planned**
Thank you

Email: r.gopalakrishnan@cris.org.in