MATLAB for Data Analytics

© 2020 The MathWorks, Inc.

MathWorks

Railway Systems



Automotive



Aeronautics



Retail



Finance



Off-highway vehicles



Industrial Automation



Oil & Gas



Clean Energy



Medical Devices



Healthcare Management



Logistics



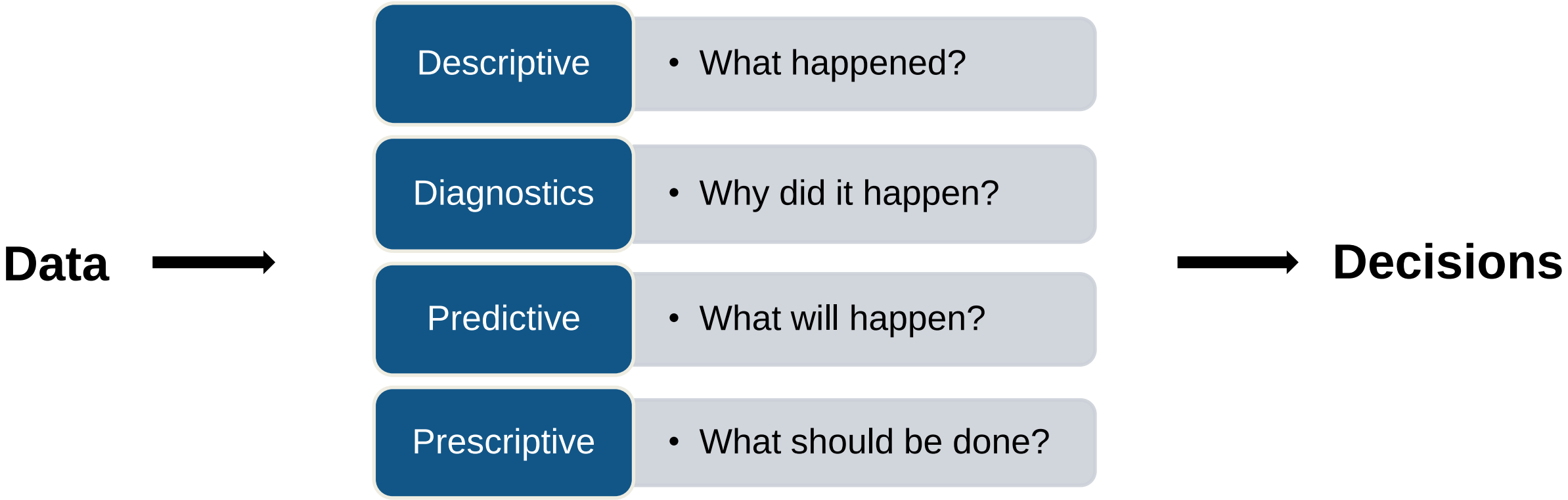
Internet



Condition Monitoring
Prognostics
Retail Analytics
Fleet Analytics
Operational Analytics
Data Analytics
Process Analytics
Risk Analysis
Health Monitoring
Supply Chain
Mfg Process Analytics
Asset Analytics
Healthcare Analytics

What is Data Analytics?

Turn large volumes of complex data into actionable information



Data Analytics Workflow

Access and Explore Data

- Files
- Databases
- Sensors

Preprocess Data

- Working with Messy Data
- Data Reduction/ Transformation
- Feature Extraction

Develop Predictive Models

- Model Creation e.g. Machine Learning
- Parameter Optimization
- Model Validation

Integrate Analytics with Systems

- Desktop Apps
- Enterprise Scale Systems
- Embedded Devices and Hardware

Data Analytics Workflow

Access and Explore Data

Files



Databases



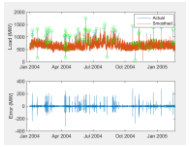
Sensors



1

Preprocess Data

Working with Messy Data



Data Reduction/ Transformation



Feature Extraction



Develop Predictive Models

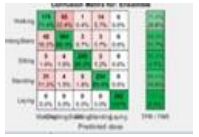
Model Creation e.g. Machine Learning



Parameter Optimization

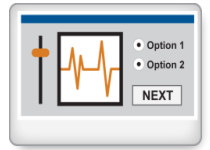


Model Validation



Integrate Analytics with Systems

Desktop Apps



Enterprise Scale Systems

MATLAB Excel .NET C/C++ .exe Java .dll

Embedded Devices and Hardware



Data Analytics Workflow

Access and Explore Data

Files



Databases



Sensors



Business and Transactional Data

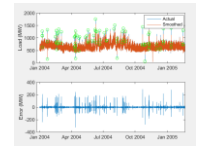
- Repositories – SQL, NoSQL, etc.
- File I/O – Text, Spreadsheet, etc.
- Web Sources – RESTful, JSON, etc.

Engineering, Scientific and Field Data

- Real-Time Sources – Sensors, GPS, etc.
- File I/O – Image, Audio, etc.
- Communication Protocols – OPC (OLE for Process Control), CAN (Controller Area Network), etc.

Preprocess Data

Working with Messy Data



Data Reduction/ Transformation



Feature Extraction



Data Analytics Workflow

Access and Explore Data

Files



Databases



Sensors

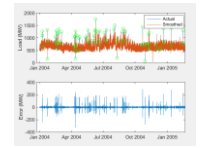


Challenges

- Data aggregation
  - Different sources (files, web, etc.)
  - Different types (images, text, audio, etc.)
- Data clean up
  - Poorly formatted files
  - Irregularly sampled data
  - Redundant data, outliers, missing data etc.
- Domain specific processing
  - Signals: Smoothing, resampling, denoising, Wavelet transforms, etc.
  - Images: Image registration, morphological filtering, deblurring, etc.
- Dealing with out of memory data (big data)

Preprocess Data

Working with Messy Data



Data Reduction/ Transformation



Feature Extraction



Data Analytics Workflow

Access and Explore Data

MATLAB Analytics work with business and engineering data 1

Preprocess Data

Files

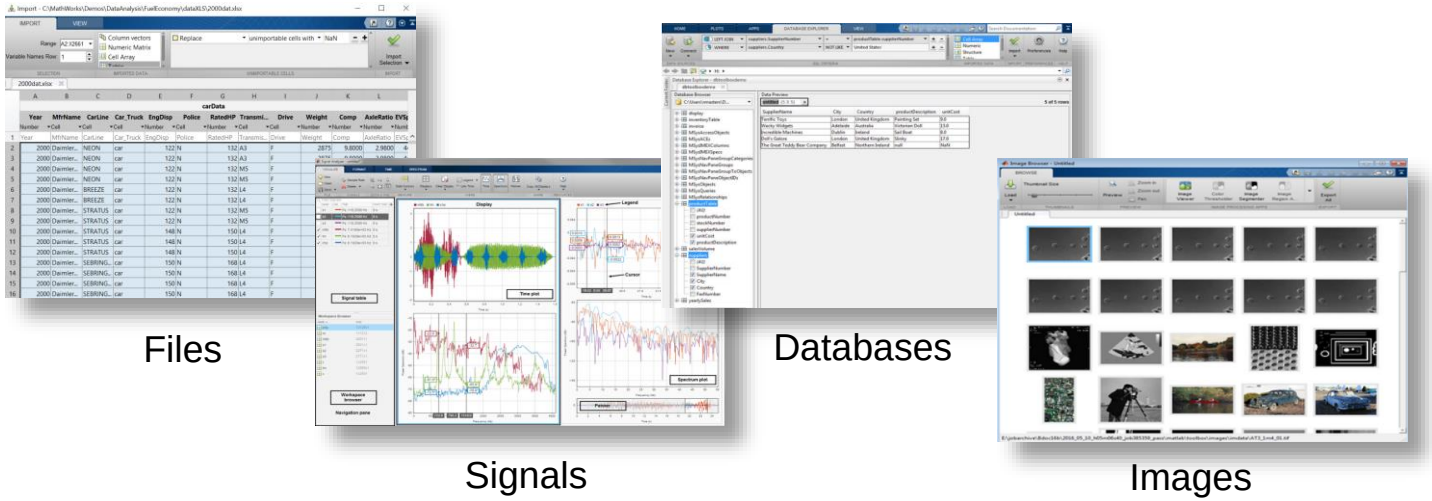


Databases



Sensors



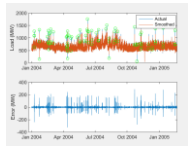


Files Databases Signals Images

- Point and click tools to access variety of data sources
- High-performance environment for big data

- Built-in algorithms for data preprocessing including sensor, image, audio, video and other real-time data

Working with Messy Data



Data Reduction/ Transformation



Feature Extraction



Data Analytics Workflow

Access and Explore Data

Files



Databases



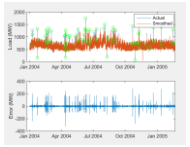
Sensors



1

Preprocess Data

Working with Messy Data



Data Reduction/ Transformation



Feature Extraction



2

Develop Predictive Models

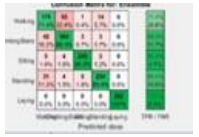
Model Creation e.g. Machine Learning



Parameter Optimization

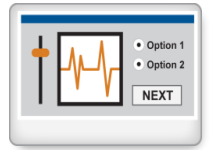


Model Validation



Integrate Analytics with Systems

Desktop Apps



Enterprise Scale Systems

MATLAB Excel .NET .exe C/C++ Java .dll

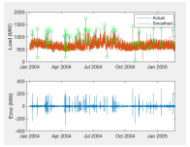
Embedded Devices and Hardware



Data Analytics Workflow

Preprocess Data

Working with Messy Data



Data Reduction/ Transformation



Feature Extraction



Challenges

- Lack of data science expertise
- Feature Extraction – How to transform data to best represent the system?
  - Requires subject matter expertise
  - No right way of designing features
- Feature Selection – What attributes or subset of data to use?
  - Entails a lot of iteration – Trial and error
  - Difficult to evaluate features
- Model Development
  - Many different models
  - Model Validation and Tuning
- Time required to conduct the analysis

Develop Predictive Models

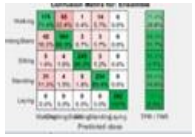
Model Creation e.g. Machine Learning



Parameter Optimization



Model Validation



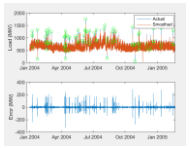
Data Analytics Workflow

Preprocess Data

MATLAB enables domain experts to do Data Science (2)

Develop Predictive Models

Working with Messy Data



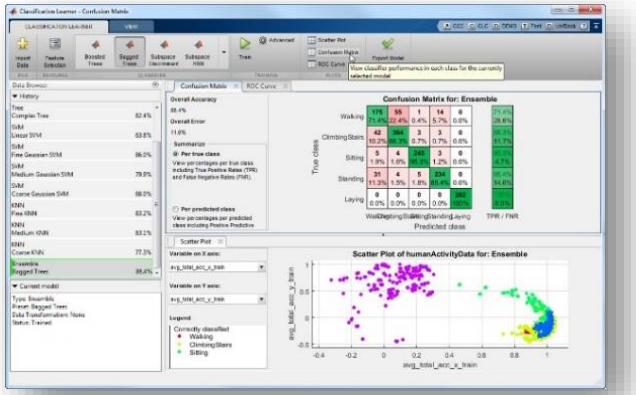
Data Reduction/ Transformation



Feature Extraction

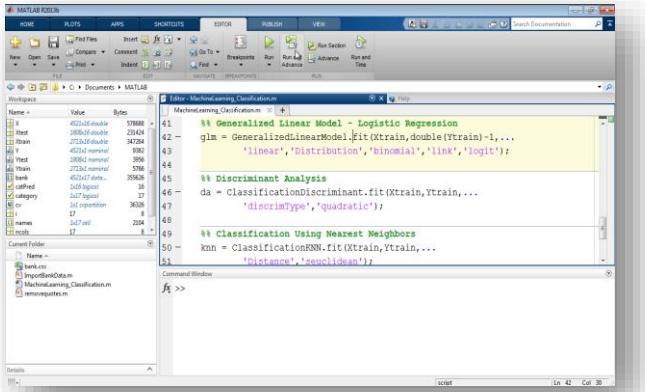


Apps



- Easy to use apps
- Wide breadth of tools to facilitate domain specific analysis
- Examples/videos to get started

Language



- Automatic MATLAB code generation
- High speed processing of large data sets

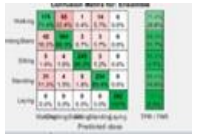
Model Creation e.g. Machine Learning



Parameter Optimization



Model Validation



Data Analytics Workflow

Access and Explore Data

Files



Databases



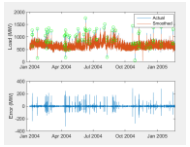
Sensors



1

Preprocess Data

Working with Messy Data



Data Reduction/ Transformation



Feature Extraction



2

Develop Predictive Models

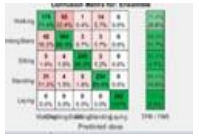
Model Creation e.g. Machine Learning



Parameter Optimization



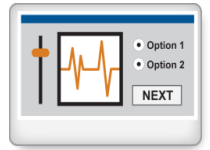
Model Validation



3

Integrate Analytics with Systems

Desktop Apps



Enterprise Scale Systems

MATLAB Excel .NET .exe C/C++ Java .dll

Embedded Devices and Hardware



Data Analytics Workflow

Develop Predictive Models

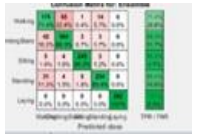
Model Creation e.g. Machine Learning



Parameter Optimization



Model Validation

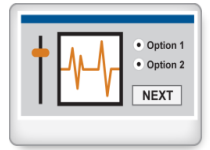


Challenges

- End user: Operators, Analysts, Administrative Staff, customers etc.
- Different target platforms:
  - Cluster or Cloud environment
  - Standalone desktop applications
  - Server based Web and enterprise systems
  - Embedded hardware
- Different Interfaces: C++, Java, Python, .NET etc.
- Need to translate analytics to production environment

Integrate Analytics with Systems

Desktop Apps



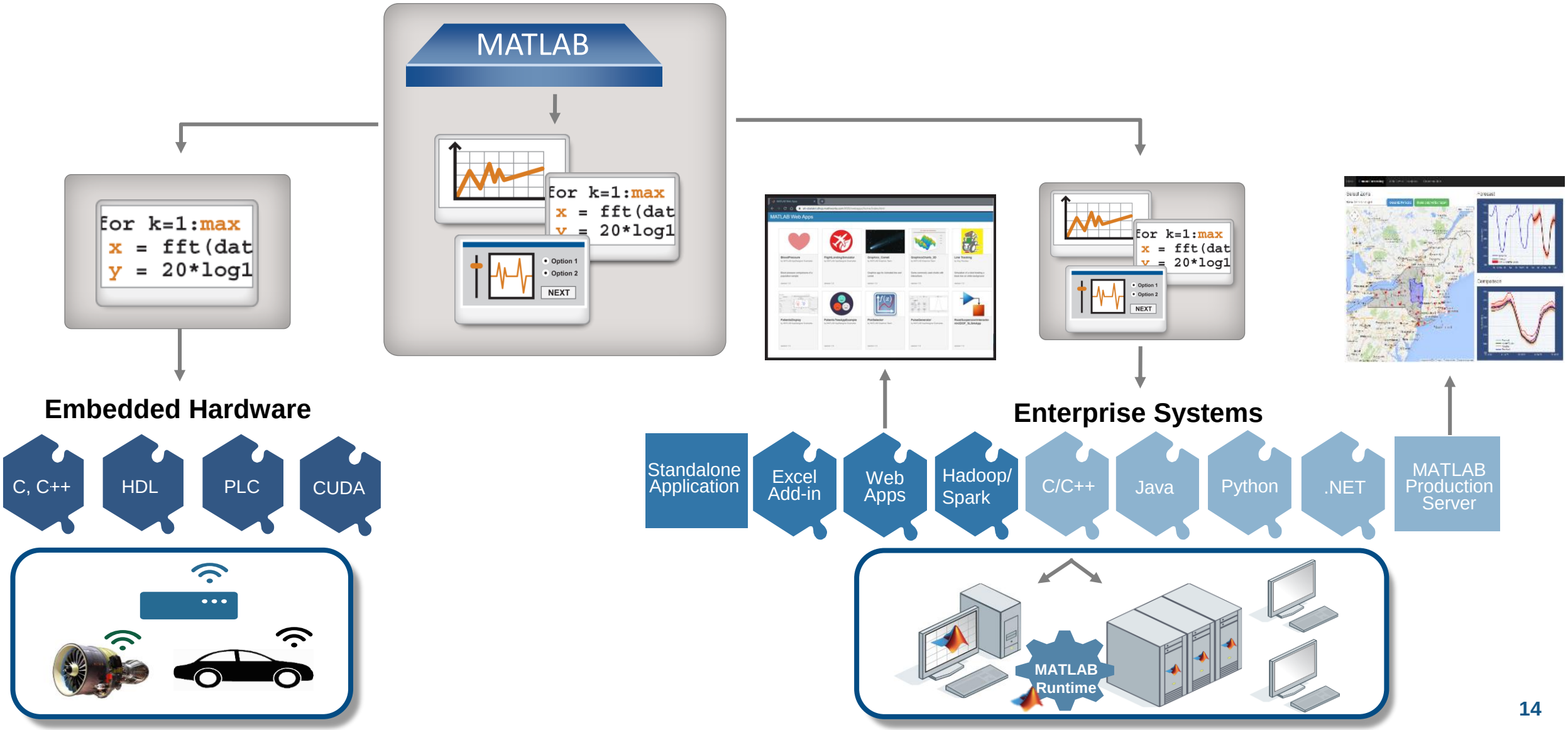
Enterprise Scale Systems

MATLAB Excel .NET .exe C/C++ Java .dll

Embedded Devices and Hardware



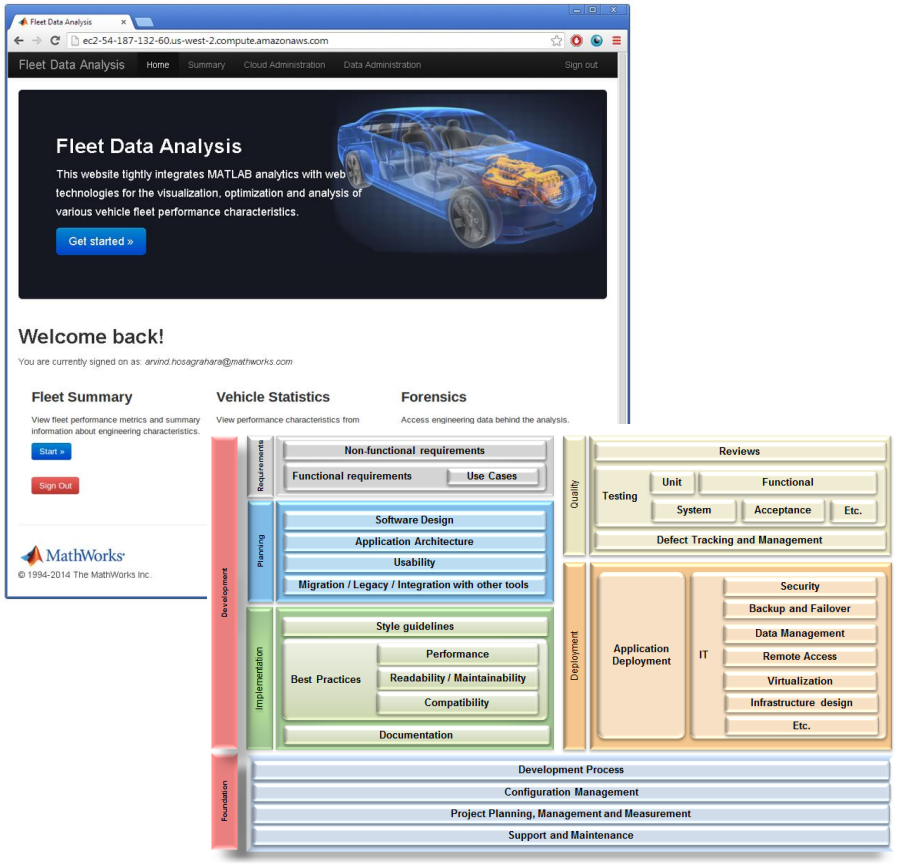
Getting out of MATLAB



Key Takeaways

1. MATLAB Analytics work with **business and engineering data**
2. MATLAB enables **domain experts to do Data Science**
3. MATLAB Analytics **run anywhere**

MathWorks Services

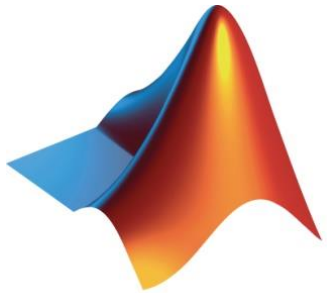


- Consulting
  - Integration
  - Data analysis/visualization
  - Unify workflows, models, data

www.mathworks.com/services/consulting/

- Training
  - Classroom, online, on-site
  - Data Processing, Visualization, Deployment, Parallel Computing

www.mathworks.com/services/training/



MathWorks®

Accelerating the pace of engineering and science

© 2020 The MathWorks, Inc. MATLAB and Simulink are registered trademarks of The MathWorks, Inc. See www.mathworks.com/trademarks for a list of additional trademarks. Other product or brand names may be trademarks or registered trademarks of their respective holders.

© 2020 The MathWorks, Inc.