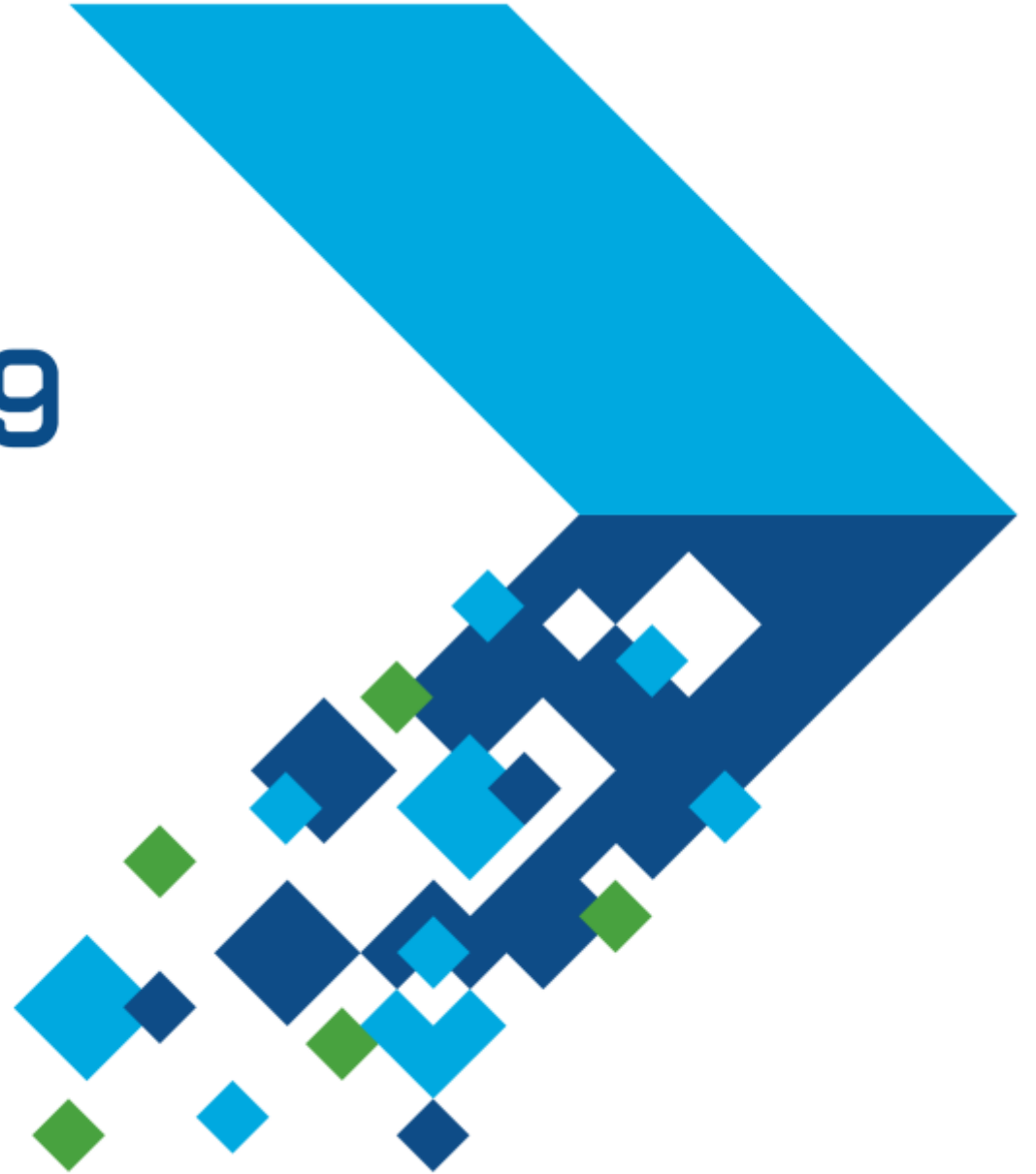


MATLAB EXPO 2019

인공지능 및 엔터프라이즈
환경 개발을 위한 빅데이터
개발 프레임워크

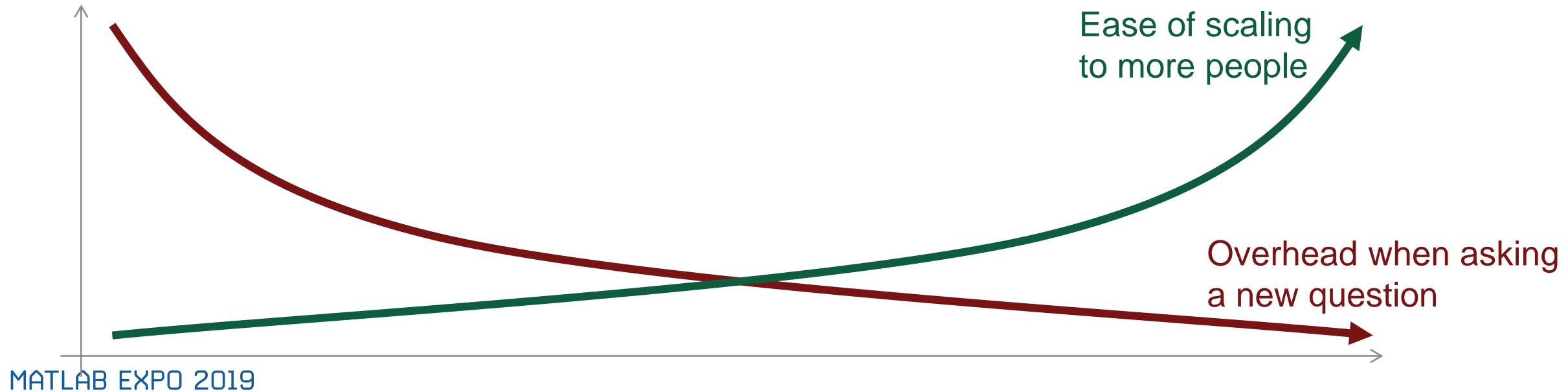
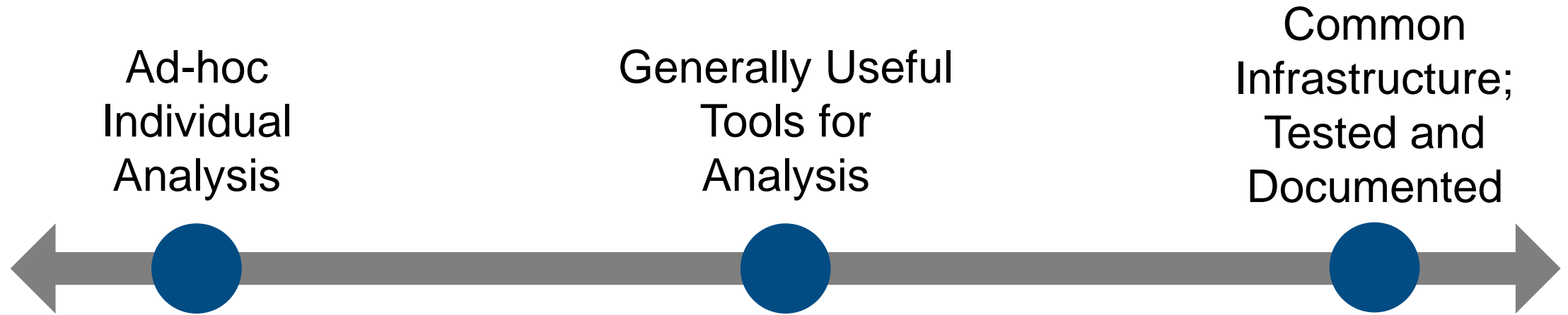
성호현



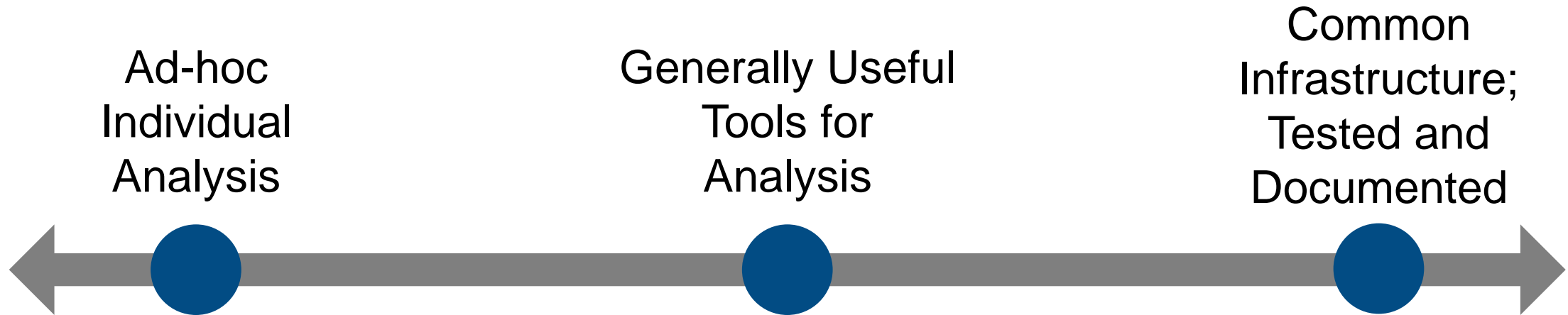
A **path** for how your team can
better work with and utilize
data.



Data Science Maturity Levels



Data Science Maturity Levels



- **Goal is to be fast: reduce time to insight**

Getting Started: Exploring a New Dataset

The image shows the MATLAB R2019a interface. The top menu bar includes 'MATLAB', 'Window', and 'Help'. The main toolbar is divided into sections: HOME, PLOTS, and APPS. The current workspace is empty. The Command Window shows the prompt `fx >>`. The Current Folder pane displays two files: `flightData_1Hz.csv` and `flightData_4Hz.csv`. The file browser shows the path `Work > MATLAB > GettingStartedWithFlightData`.

Workspace

| Name | Value |
|------|-------|
|------|-------|

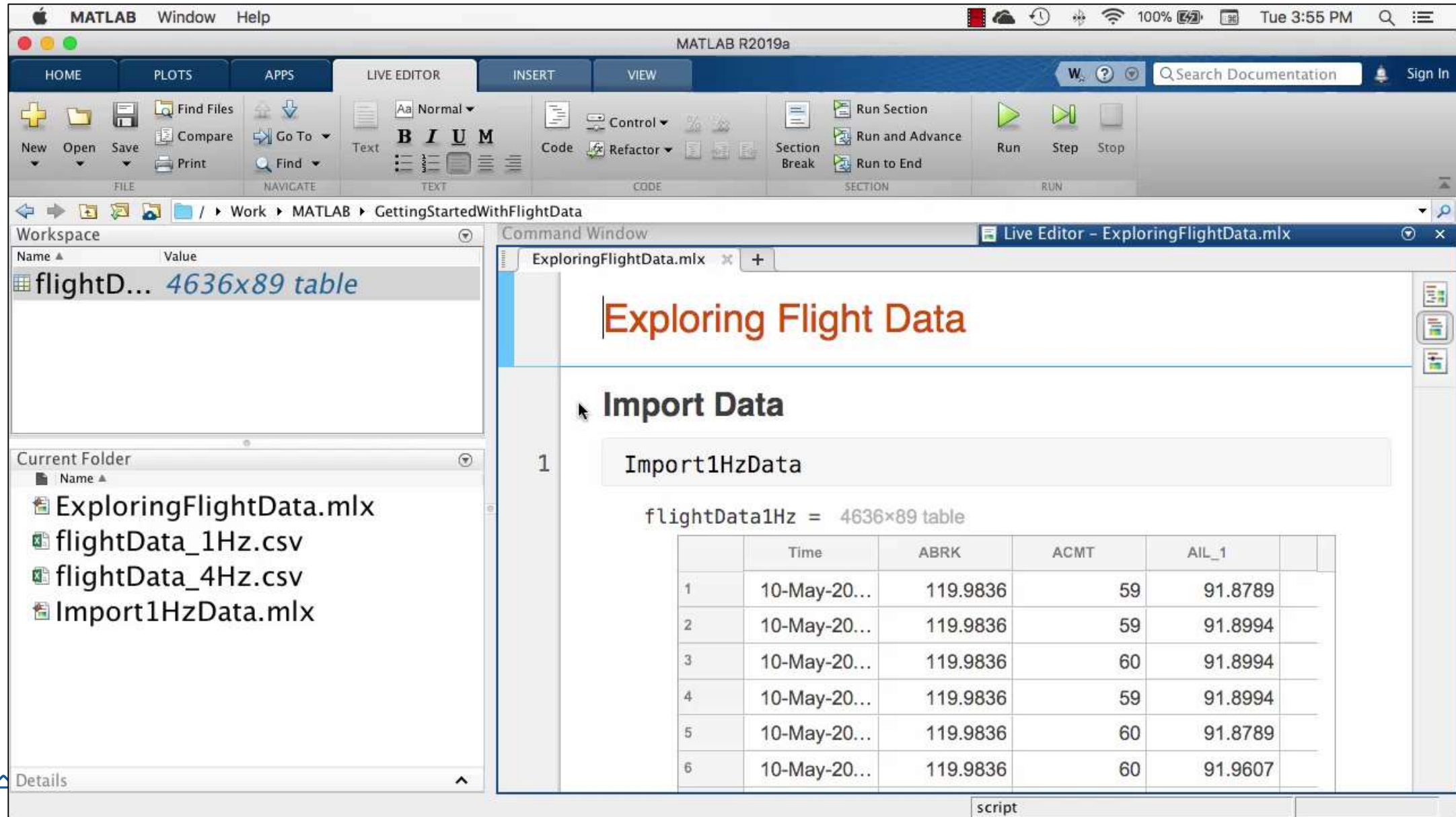
Current Folder

- flightData_1Hz.csv
- flightData_4Hz.csv

Command Window

```
fx >>
```

Getting Started: Exploring a New Dataset



The screenshot shows the MATLAB R2019a interface. The workspace contains a variable named `flightD...` with a value of `4636x89 table`. The current folder contains the following files:

- ExploringFlightData.mlx
- flightData_1Hz.csv
- flightData_4Hz.csv
- Import1HzData.mlx

The live editor window displays the title "Exploring Flight Data" and the section "Import Data". The code in the live editor is:

```
1 Import1HzData
```

The output of the code is a table with 6 rows and 5 columns:

| | Time | ABRK | ACMT | AIL_1 |
|---|--------------|----------|------|---------|
| 1 | 10-May-20... | 119.9836 | 59 | 91.8789 |
| 2 | 10-May-20... | 119.9836 | 59 | 91.8994 |
| 3 | 10-May-20... | 119.9836 | 60 | 91.8994 |
| 4 | 10-May-20... | 119.9836 | 59 | 91.8994 |
| 5 | 10-May-20... | 119.9836 | 60 | 91.8789 |
| 6 | 10-May-20... | 119.9836 | 60 | 91.9607 |

Getting Started: Exploring a New Dataset

The screenshot shows the MATLAB R2019a interface. The workspace contains two variables: `ans` (4636x88 double) and `flightD...` (4636x89 table). The current folder contains `ExploringFlightData.mlx`, `flightData_1Hz.csv`, `flightData_4Hz.csv`, and `Import1HzData.mlx`. The live editor window shows the following code and output:

```
4 flightData1Hz.ABRK
```

Access variables by type:

```
5 flightData1Hz{: , vartype("numeric")}]
```

```
ans = 4636x88
106 x
    0.0001    0.0001    0.0001    0.0001    0.0060 ...
    0.0001    0.0001    0.0001    0.0001    0.0060
    0.0001    0.0001    0.0001    0.0001    0.0060
    0.0001    0.0001    0.0001    0.0001    0.0060
    0.0001    0.0001    0.0001    0.0001    0.0060
    0.0001    0.0001    0.0001    0.0001    0.0060
    0.0001    0.0001    0.0001    0.0001    0.0060
    0.0001    0.0001    0.0001    0.0001    0.0060
    0.0001    0.0001    0.0001    0.0001    0.0060
    0.0001    0.0001    0.0001    0.0001    0.0060
```

script Ln 5 Col 36

Getting Started: Exploring a New Dataset

Workspace

| Name | Value |
|------------|----------------------|
| ans | 1x1 StackedLineChart |
| endtime | 1x1 datetime |
| flightD... | 4636x89 table |
| starttime | 1x1 datetime |

Current Folder

- ExploringFlightData.mlx
- flightData_1Hz.csv
- flightData_4Hz.csv
- Import1HzData.mlx

Command Window

ExploringFlightData.mlx * * +

off_1
off_2
off_3
off_4

Time

script Ln 12 Col 21

Missing Data
`ismissing`
`rmissing`
`fillmissing`

Outliers
`isoutlier`
`rmoutliers`
`filloutliers`

Change Points
`ischange`

Noisy Data
`smoothdata`

and more...

<https://www.mathworks.com/help/matlab/preprocessing-data.html>

Getting Started: Exploring a New Dataset

The image shows the MATLAB R2019a interface. The top menu bar includes 'MATLAB', 'Window', and 'Help'. The main toolbar has tabs for 'HOME', 'PLOTS', 'APPS', 'LIVE EDITOR', 'INSERT', and 'VIEW'. Below the toolbar are icons for file operations (New, Open, Save, Find Files, Compare, Print), navigation (Go To, Find), text formatting (Normal, Bold, Italic, Underline, Monospace), code editing (Control, Refactor), and execution (Run Section, Run and Advance, Run to End, Run, Step, Stop).

The workspace on the left shows the following variables:

| Name | Value |
|------------|----------------------|
| ans | 1x1 StackedLineChart |
| endtime | 1x1 datetime |
| flightD... | 4636x89 table |
| starttime | 1x1 datetime |

The current folder on the left contains the following files:

- ExploringFlightData.mlx
- flightData_1Hz.csv
- flightData_4Hz.csv
- Import1HzData.mlx
- Import4HzData.mlx

The command window on the right shows the following code:

```

15 Import4HzData
16 t4hz = table2timetable(flightData4Hz)

```

The plot window shows a line graph of flight data. The x-axis is labeled 'Time' and ranges from 17:15 to 18:15 on May 10, 2001. The y-axis ranges from 60 to 90. The plot shows a single data point at approximately 17:15 with a value of about 85.

Load 4Hz Data
Load data from same flight that was sampled at 4Hz.

Synchronize 1Hz and 4Hz Data
Join the 1Hz and 4Hz data using the synchronize command. The synchronize command gives us flexibility in how the synchronize occurs. Here, we use the default synchronize method which synchronizes the

Getting Started: Exploring a New Dataset

The MATLAB R2019a interface is shown with the following components:

- Workspace:** Lists variables including 'ans' (1x1 StackedLineChart), 'endtime' (1x1 datetime), and two 'flightD...' tables (4636x89 and 18544x50).
- Current Folder:** Shows files like 'ExploringFlightData.mlx', 'flightData_1Hz.csv', 'flightData_4Hz.csv', 'Import1HzData.mlx', and 'Import4HzData.mlx'.
- Command Window:** Displays a table of flight data:

| Index | Date | Value |
|-------|--------------|----------|
| 3 | 10-May-20... | 119.9836 |
| 4 | 10-May-20... | 119.9836 |
| 5 | 10-May-20... | 119.9836 |
| 6 | 10-May-20... | 119.9836 |
| 7 | 10-May-20... | 119.9836 |
| 8 | 10-May-20... | 119.9836 |
| 9 | 10-May-20... | 119.9836 |
- Live Editor:** Shows a script with the following code:


```

19 figure;
20 inFlight = t.WOW==1;
21 geoplot(t.LATP(inFlight),t.LONP(inFlight),'Linewidth',3);
      
```

Four different ways to visualize geographic data are shown:

- geoplot:** A line plot showing a flight path on a map of the Pacific Ocean, with latitude on the y-axis (39 to 42) and longitude on the x-axis (-86 to -82).
- geoscatter:** A scatter plot showing data points on a world map, with latitude on the y-axis (-60 to 60) and longitude on the x-axis (-100 to 100).
- geobubble:** A bubble plot showing data points on a world map, with latitude on the y-axis (30°N to 50°N) and longitude on the x-axis (20°W to 50°W).
- geodensityplot:** A density plot showing data points on a world map, with latitude on the y-axis (10 to 60) and longitude on the x-axis (100 to 200).

Visualizing the Flight Path

<https://www.mathworks.com/help/matlab/geographic-plots.html>

Data Science Maturity Levels



- **Explore and understand data**
- **Document analysis**
- **Tools will be re-used in next steps**

Data Science Maturity Levels






- **Apply to different datasets**
 - **Functions/Scripts**
 - **MATLAB Apps**
- **Trend: Work with **BIG DATA****

Overview of Flight Data

- **35** unique aircraft
- **180,000** unique flights
- **300 GB** of data
- Source:
 - NASA Dash Link: Sample Flight Data
 - <https://c3.nasa.gov/dashlink/projects/85/>



Big Data Creates Opportunities

-  Find rare events, then dive deeper
-  Build and validate test scenarios that match real-world conditions
-  Perform fleet-wide calculations

Big Data Requires New Tools

Create a datastore from all CSV files

```
ds = datastore('*.*.csv')
```

Read a single file of data

```
data = read(ds);
```

Reset the datastore back to the first file

```
reset(ds);
```

Find the maximum value of “Y” in each file

```
X = [];
while hasdata(ds)
    data = read(ds);
    X(end+1) = max(data.Y);
end
```

| Built-In Datastores | |
|------------------------|--------------------------------|
| General | datastore |
| | spreadsheetDatastore |
| | tabularTextDatastore |
| | fileDatastore |
| Database | databaseDatastore |
| Image | imageDatastore |
| | denoisingImageDatastore |
| | randomPatchExtractionDatastore |
| | pixelLabelDatastore |
| | augmentedImageDatastore |
| Audio | audioDatastore |
| Predictive Maintenance | fileEnsembleDatastore |
| | simulationEnsembleDatastore |
| Simulink | SimulationDatastore |
| Automotive | mdfDatastore |

Big Data Requires New Tools

Custom Datastore

- Customize a datastore to work with your dataset
- Gives you control over how data is loaded and formatted
- MATLAB subclass: “fill-in-the-blanks”
- Build a piece of infrastructure, then re-use it in your analyses

```
function [data,info] = read(ds)
    ...
end
```

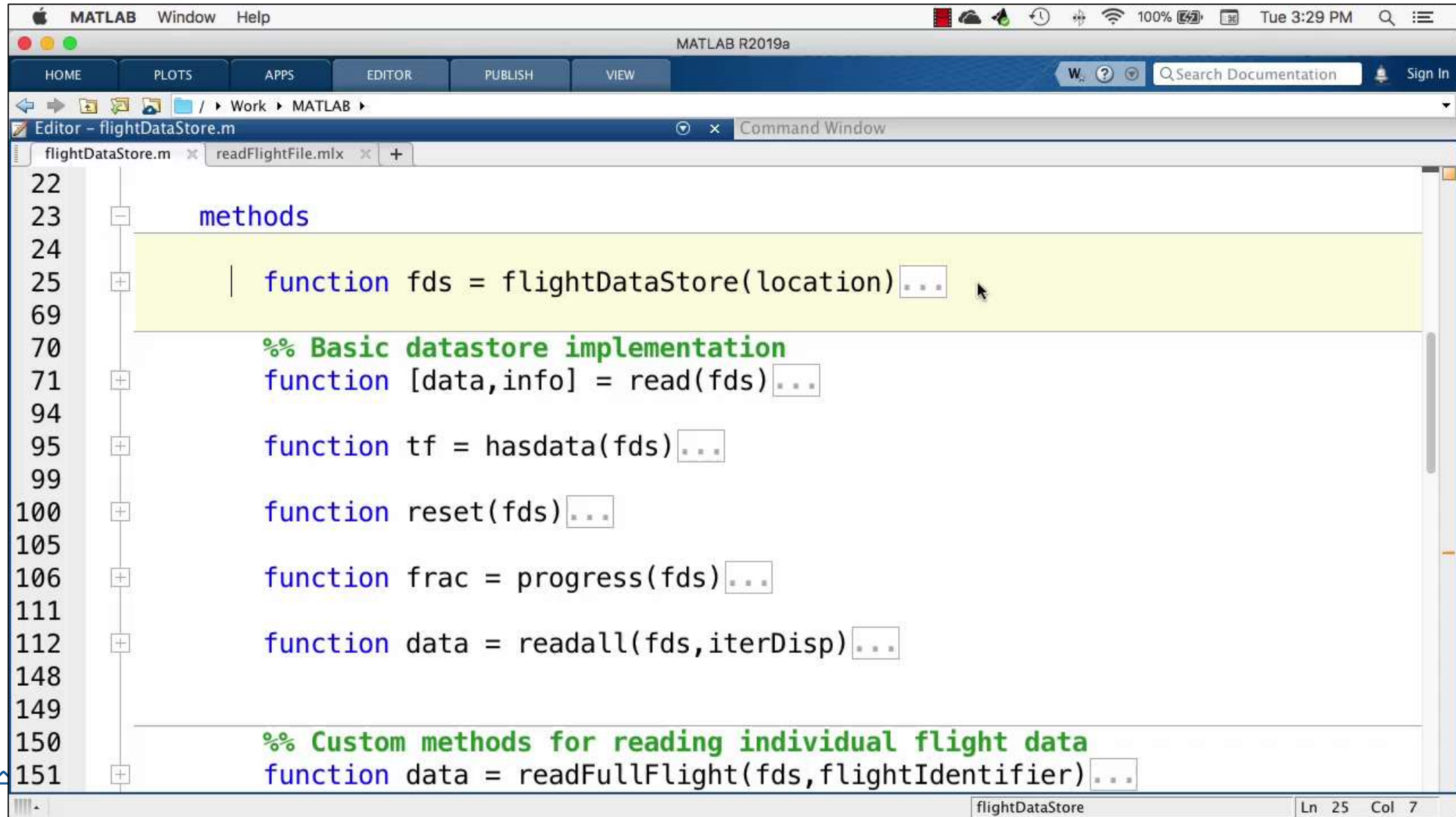
```
function tf = hasdata(ds)
    ...
end
```

```
function reset(ds)
    ...
end
```

```
function p = progress(ds)
    ...
end
```

```
function data = readall(ds)
    ...
end
```


A Custom Datastore for Flight Data



The image shows a MATLAB R2019a window with the Editor and Command Window. The Editor displays the code for a custom datastore named 'flightDataStore.m'. The code is organized into sections with expandable/collapsible icons on the left margin.

```
22  
23 methods  
24  
25 | function fds = flightDataStore(location) ...  
69  
70 %% Basic datastore implementation  
71 | function [data,info] = read(fds) ...  
94  
95 | function tf = hasdata(fds) ...  
99  
100 | function reset(fds) ...  
105  
106 | function frac = progress(fds) ...  
111  
112 | function data = readall(fds,iterDisp) ...  
148  
149  
150 %% Custom methods for reading individual flight data  
151 | function data = readFullFlight(fds,flightIdentifier) ...
```

The status bar at the bottom indicates the current file is 'flightDataStore' and the cursor is at line 25, column 7.



Find Rare Events, then Dive Deeper

The image shows the MATLAB R2019a interface. The top menu bar includes 'MATLAB', 'Window', and 'Help'. The title bar indicates 'MATLAB R2019a'. The main window is titled 'Live Editor - IncidentAnalysisExample.mlx' and contains the following content:

Incident Analysis

The data for tail 666 captures an engine fire that happened shortly after takeoff. This example analyzes data from that tail number to see if there is any interesting information that can be learned about how/when the fire occurred.

Search for Flights where the Engine 2 Fire Sensor Went True

Loop through all flights for this aircraft, and display some information if there is a fire.

```
1 ds = flightDataStore('/Volumes/Ext/Data/FlightData/666/');|
2 ds.SelectedVariableNames = {'FIRE_2'};
3 ds.AddFlightIdentifier = true;

4 results = table('Size',[0 2],'VariableTypes',{'categorical','double'},...
5     'VariableNames',{'Flight','Fire_Seconds'});
6 while hasdata(ds)
7     t = read(ds);
8     if any(t.FIRE_2)
9         results = [results; {t.Flight(1), sum(t.FIRE_2)}];
10    end
11 end
12 results
```

The status bar at the bottom shows 'script' and 'Ln 1 Col 59'.

Perform Fleet-Wide Calculations

The image shows the MATLAB R2019a interface. The top menu bar includes 'HOME', 'PLOTS', 'APPS', 'LIVE EDITOR', 'INSERT', and 'VIEW'. The 'LIVE EDITOR' tab is active, showing a script titled 'FleetWideFuelEconomy.mlx'. The script content is as follows:

Fleet-Wide Calculations

Calculate fuel efficiency across all of the flights in the dataset.

Tall Arrays

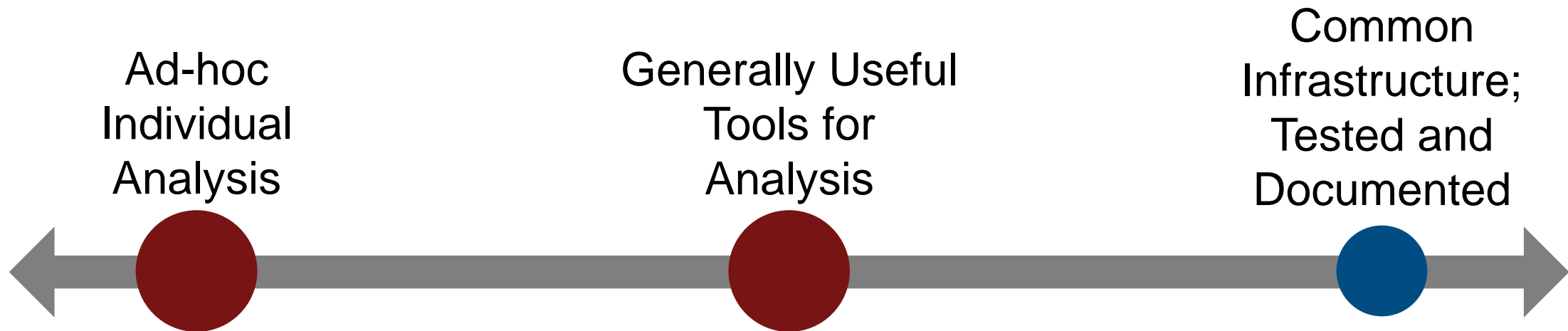
Creating a tall array from the datastore gives us access to many built-in functions.

```
1 ds = flightDataStore('/Volumes/Ext/Data/FlightData/');
```

TAS: True Airspeed (knots)
 FF_N: Fuel Flow N (lbs/hr)
 WOW: Weight On Wheels (logical)

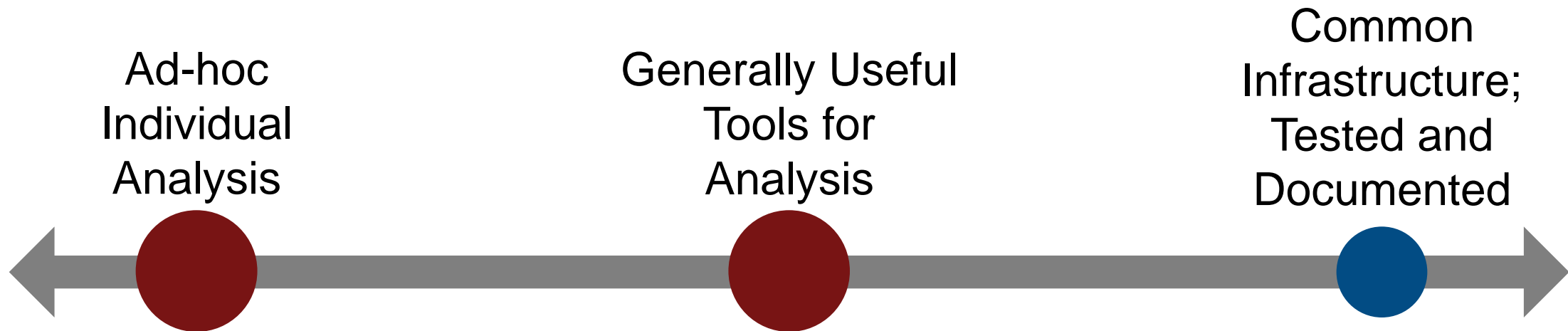
The Command Window at the bottom shows the output of the script: 'ds SelectedVariableNames [TAS] FF_N FF_N FF_N FF_N FF_N [WOW]'. The status bar at the bottom right indicates 'script'.

Data Science Maturity Levels



- **Make it easy to navigate the data**
- **Re-use each time you analyze the dataset**

Data Science Maturity Levels



- **Collaborate: Work with others on a common code base**
- **Verify: Write well-tested software**
- **Share: Build tools for others**

MATLAB Projects

The screenshot displays the MATLAB R2019a desktop environment. The top menu bar includes 'MATLAB', 'Window', and 'Help'. The main toolbar is organized into sections: HOME, PLOTS, and APPS. The HOME section contains icons for 'New Script', 'New Live Script', 'New', 'Open', 'Compare', 'Import Data', 'Save Workspace', 'Open Variable', and 'Clear Workspace'. The APPS section includes 'Analyze Code', 'Run and Time', 'Clear Commands', 'Layout', 'Parallel', 'Add-Ons', 'Help', 'Community', 'Request Support', and 'Learn MATLAB'. The Command Window is open, showing the prompt 'fx >>' and a mouse cursor. The window title is 'Project - FlightDataToolbox'. The system tray at the top right shows the date and time as 'Mon 1:51 PM'.

Testing

The screenshot shows the MATLAB R2019a Project - FlightDataToolbox interface. The top toolbar includes sections for FILE, TOOLS, ENVIRONMENT, and SOURCE CONTROL. The main area displays a file browser view of the project files.

| Name | Status | Git | Size | Date Modified |
|-------------------------|--------|-----|--------|----------------|
| +tests | ✓ | · | | 2/28/19, 1:... |
| doc | ✓ | · | | 3/1/19, 11:... |
| flightDataStore.m | ✓ | ● | 13 KB | 3/4/19, 11:... |
| FlightDataToolbox.mltbx | ✓ | ● | 426 KB | 3/1/19, 11:... |
| info.xml | ✓ | ● | 2 KB | 3/1/19, 9:5... |
| PackageToolbox.prj | ✓ | ● | 6 KB | 3/1/19, 11:... |
| readFlightFile.mlx | ✓ | ● | 5 KB | 2/26/19, 10... |

The Command Window at the bottom shows the prompt `fx >>`.

Creating a Toolbox

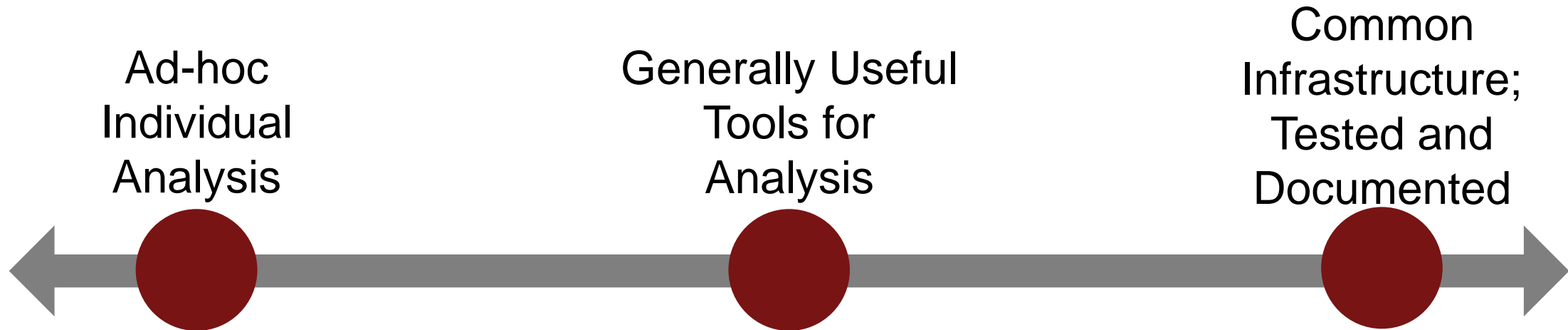
The screenshot shows the MATLAB R2019a interface. The top menu bar includes 'HOME', 'PLOTS', 'APPS', 'PROJECT', and 'PROJECT SHORTCUTS'. The 'PROJECT SHORTCUTS' tab is active, showing icons for 'New', 'Open', 'Add Files', 'Unsaved Changes', 'Share', 'Search', 'Custom Tasks', 'Run Checks', 'References', 'Details', 'Project Path', 'Startup Shutdown', 'Git Details', 'Refresh', 'Commit', 'Fetch', 'Push', 'Pull', 'Remote', 'Branches', 'Submodules', and 'Stashes'. The Command Window displays the project structure for 'FlightDataToolbox'.

Command Window: Project - FlightDataToolbox

| Name | Status | Git | Size | Date Modified |
|-------------------------|--------|-----|--------|----------------|
| +tests | ✓ | · | | 2/28/19, 1:... |
| doc | ✓ | · | | 3/1/19, 11:... |
| flightDataStore.m | ✓ | ● | 13 KB | 3/4/19, 11:... |
| FlightDataToolbox.mltbx | ✓ | ● | 426 KB | 3/4/19, 2:0... |
| info.xml | ✓ | ● | 2 KB | 3/1/19, 9:5... |
| PackageToolbox.prj | ✓ | ● | 6 KB | 3/1/19, 11:... |
| readFlightFile.mlx | ✓ | ● | 5 KB | 2/26/19, 10... |

Git status: Current branch: master, Branch status: Normal, Ahead of /origin/master

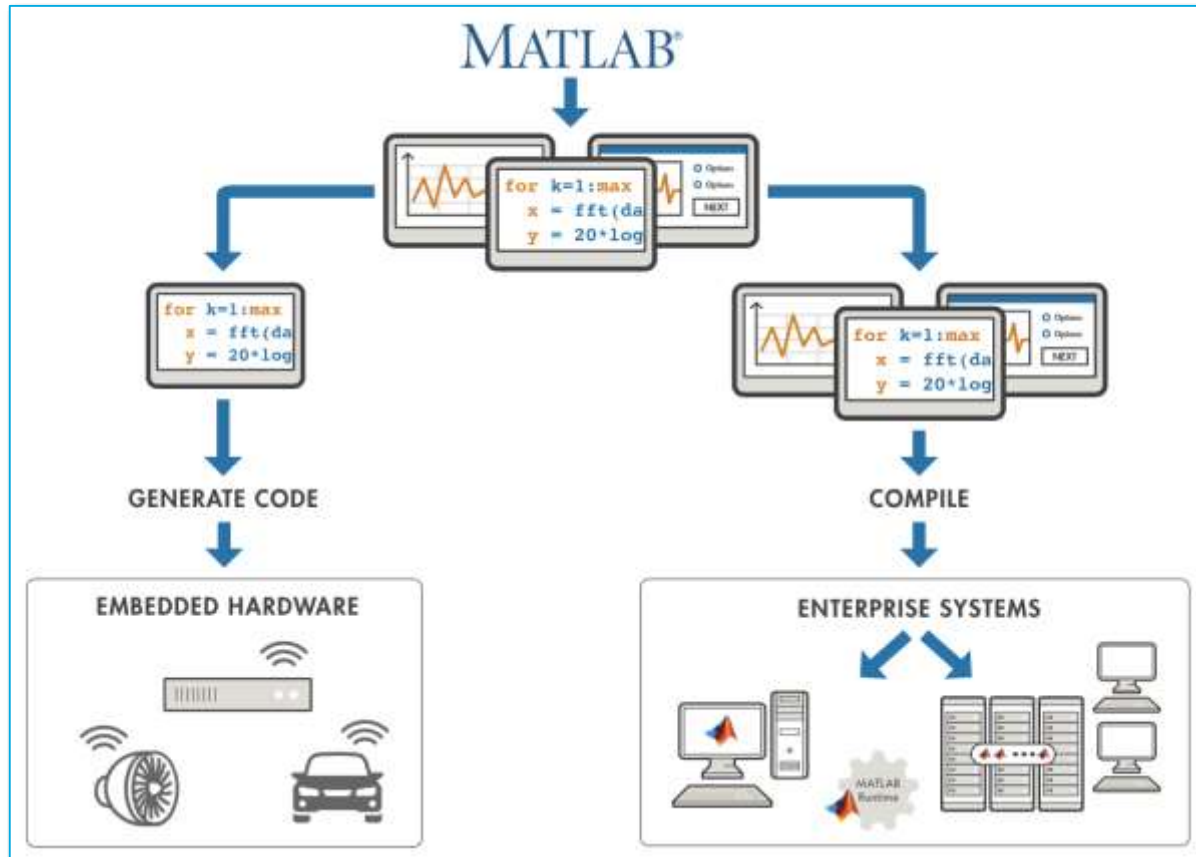
Data Science Maturity Levels



- **Scale-out to larger group of users**
- **Easier to maintain and share**

What's Next?

Advanced Analytics and Machine Learning



Build and Test Algorithms for Embedded Systems

Deploy Apps and Analytics to Enterprise IT Systems

Takeaways

- MATLAB has many new tools to help you **better work with and utilize your data**
- Create tools for **you / your team / your organization** to explore and analyze data
- Increasing maturity with data science is a journey; we're here to help

MATLAB EXPO 2019

데모 부스와 상담부스로 질문 하시기 바랍니다.

감사합니다

