



Ubiquitous Data Analysis ---- Construct Platform for Enterprise

Gang Wu



MATLAB **EXPO** 2021

Who we are

Beijing Innovation Center, P&G
focusing on fast consumable goods and devices



What we do



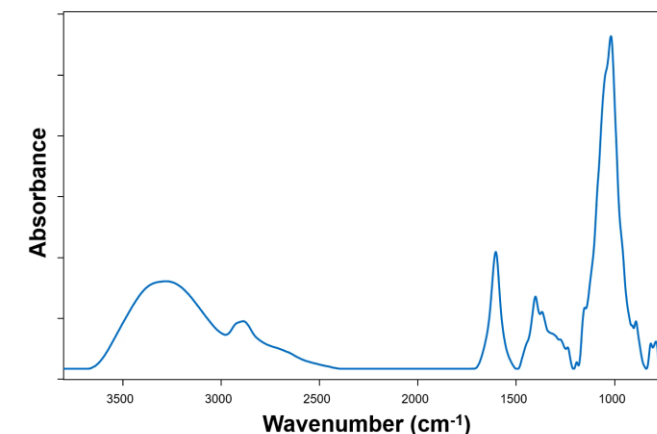
- ❑ To Innovate Product from:
 - Consumer Preference
 - Formulation & Process
 - Raw Material
 - Safety
 - Efficacy
 - Quality

- ❑ By multi-function measurement:
 - Sensorial
 - Appearance
 - Microbiological
 - Clinical
 - Analytical
 - Chemical
 - Physical
 - Instrumental
 - Imaging

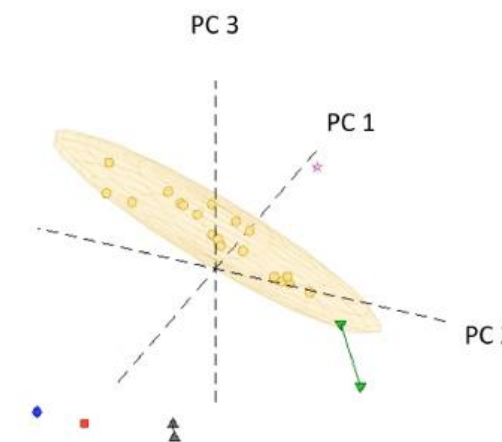
What we use



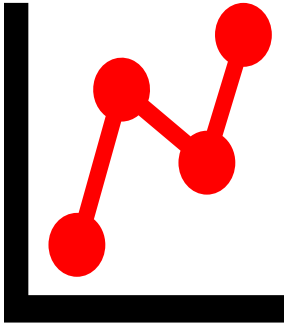
Business Insight



Visualization

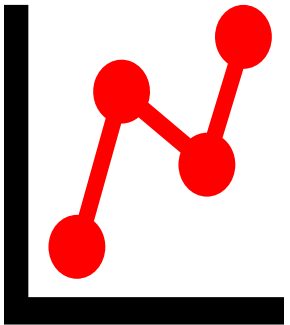


My real case



- To build understanding of formulation & flavor interaction to improve microbiological hostility of personal cleaning product
- The study started by flavor library & formulae space screening
- 1st round, flavor & formulation understanding go deeper, so DOE was introduced, like: flavor A > flavor B
- 2nd round, more flavor & formulation add, breaks down into chemical ingredients, like: ingredient A > ingredient B

My real case



- 3rd round, spectroscopy measurement (chemical information) was added, but prove not sensitive enough.
- 4th round, change to another spectroscopy measurement, find good relationship, but sample size not big enough
- 5th round, more sample data collected, but forget to collect data for “internal standard”, which calibrate offset of microbiological data from each round
- 6th round, more sample & “internal standard” data collected, the model built up successfully

What I need

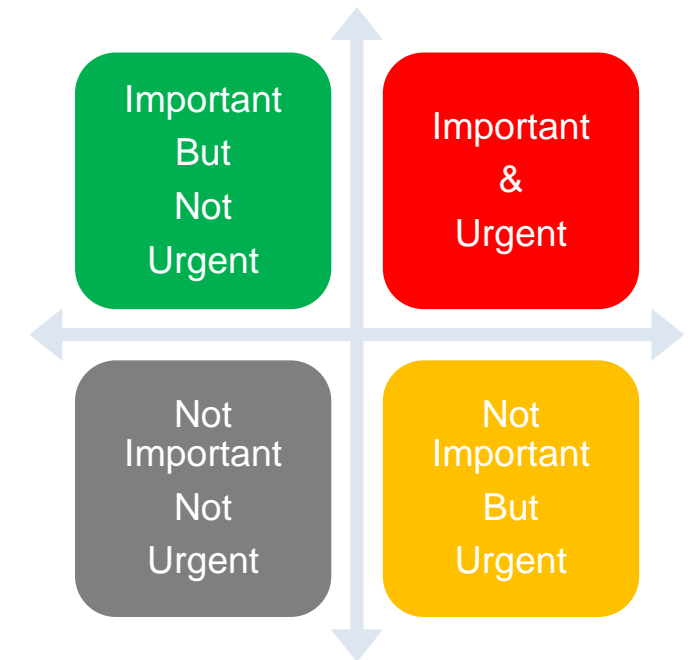
Goal: Serve business partners faster, better, cheaper

Growth Mindset:

- 80/20 principle is fact
- I wish to **ONLY** focus on the most important
- I am not productive enough yet

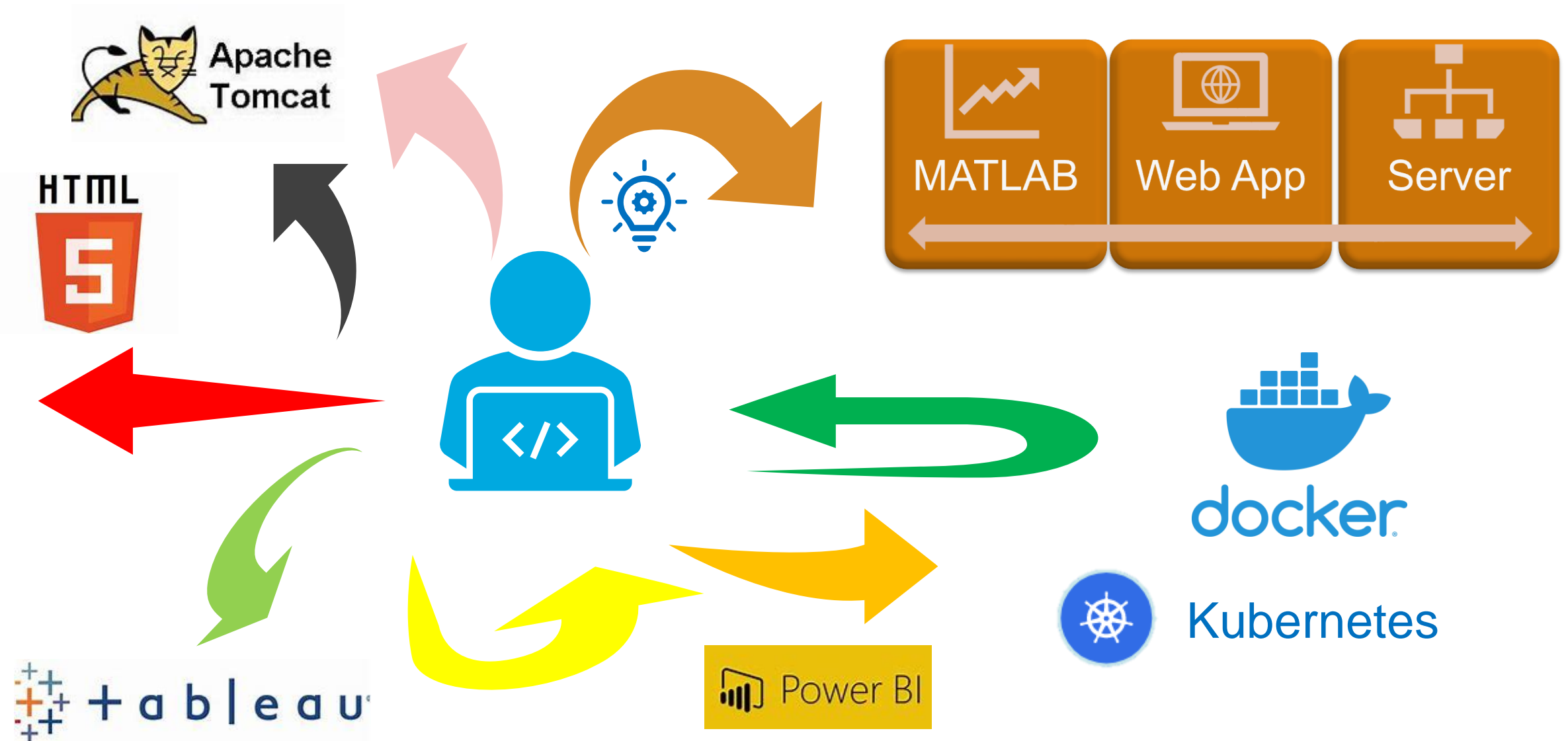
Solution:

- ~~Recruit someone work for me~~
- ~~Ideal job assignment~~
- ~~Train my customer with basic skills~~
- Improve productivity of routine tasks
- with minimal cost, timing & effort
- for both data analyst & business partners



授人与鱼，
不如授人以 ~~渔~~
更不如授人以外卖！

What I found



MATLAB

- Stats & Machine Learning
- Deep Learning
- Parallel Computing
- Image Acquisition
- Image Process
- Computer Vision
- Signal Process

3rd Parties

- M model
- M function
- CUDA
- C++

App Designer



Compiler

Web App Server

- App-1
- App-2
-

Web Browser



Intranet

Admin

- Admin-1
- Admin-2

User

- User-1
- User-2

Group

- Group-1
- Group-2

My Journey



Azure Active Directory

Virtual Machine with resizable

- CPU
- GPU
- Storage

Operation System

- Remote Access
- Fix IP address
- Computer Name & DNS
- SSL certificate

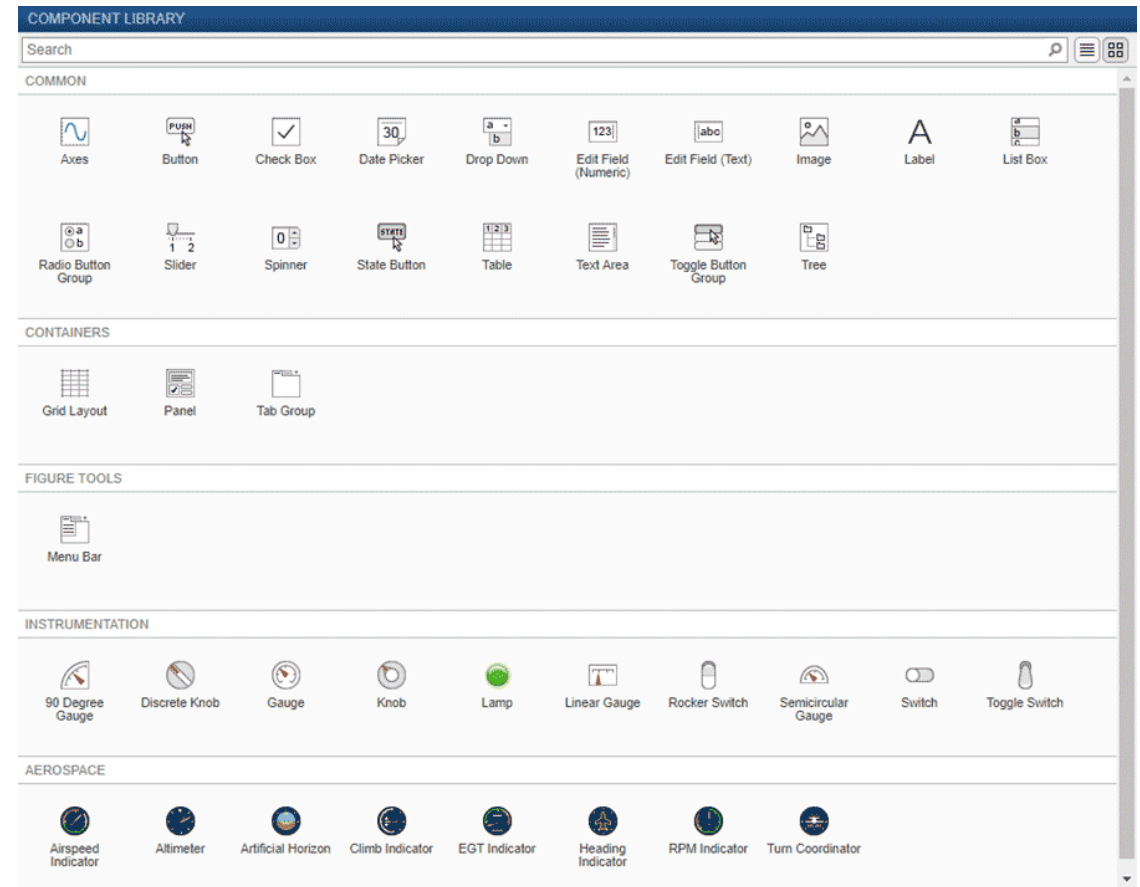
Domain Controller

- User Access & Group
- Administrator Privilege
- Share Drive to map

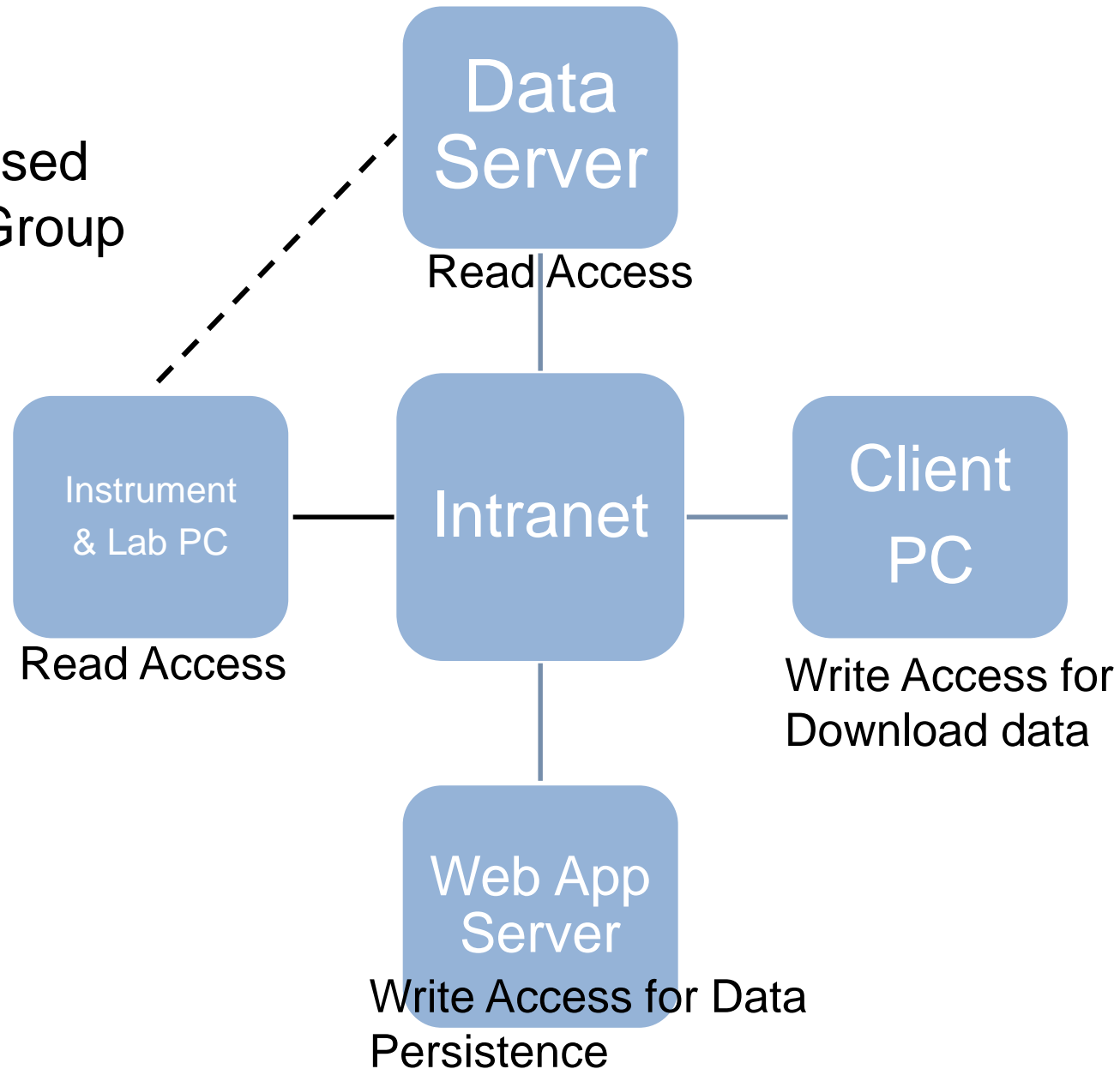
MATLAB Web App Server / services

My Apps

- Graphical user interface (GUI)
- GUIDE could migrate
- Drag and with visual components
- Auto create the object-oriented code
- Code Analyzer give warning
- Rich components and controls
- Components allow customization
- Edit the callbacks
- Auto publish to web app server
- No worry about version control
- No need runtime installation



Prepare Domain based
Special ID and/or Group
for user access
& data transfer



My goal



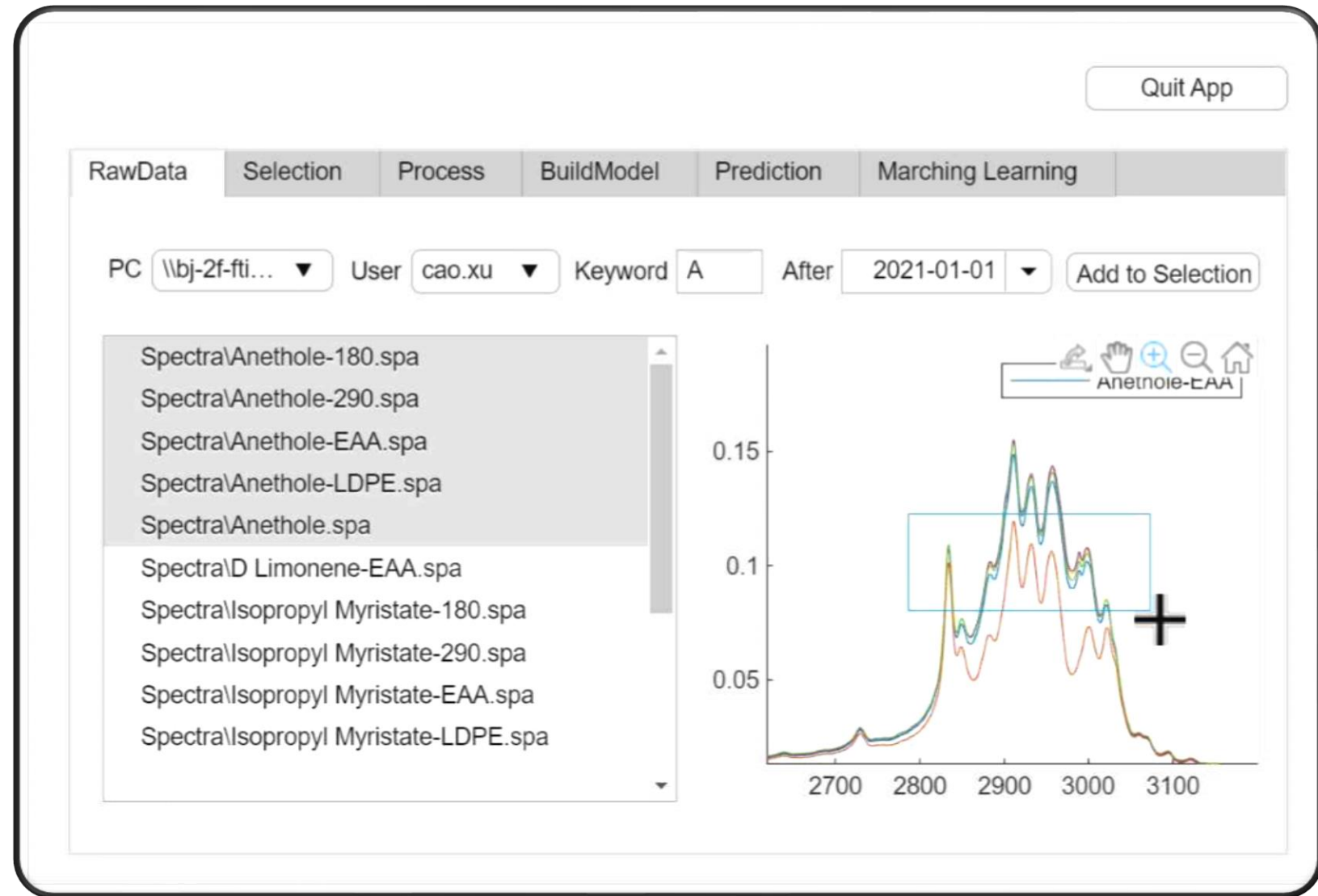
- Spectroscopy library of formula & flavor ingredients pre-scanned & stored on instrument PC
- Partner use existing spectroscopy of ingredients as X, available historical microbiological hostility data as Y, to build model via web app, several rounds
- Non-reliable results by web app indicate data not enough
Poor results by web app indicate poor correlation

My goal

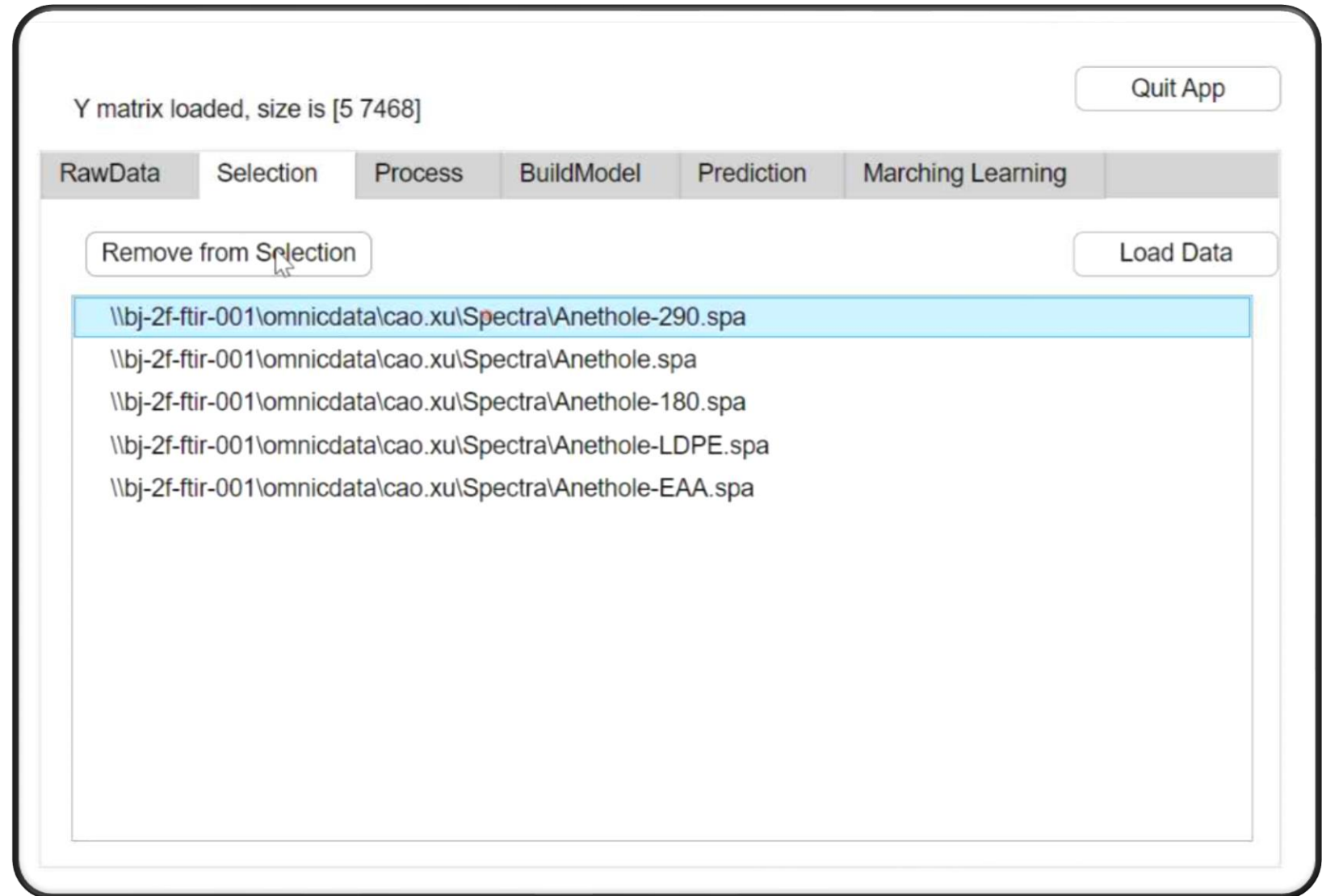


- Partner will collect more data if non-reliable
- Build model by web app, or reach me for final tune
- Use model to predict those existing flavor without microbiological hostility data, top candidates go to small scale pilot production
- Test microbiological hostility by new production, if good then project succeed, if bad then more data for next round modelling

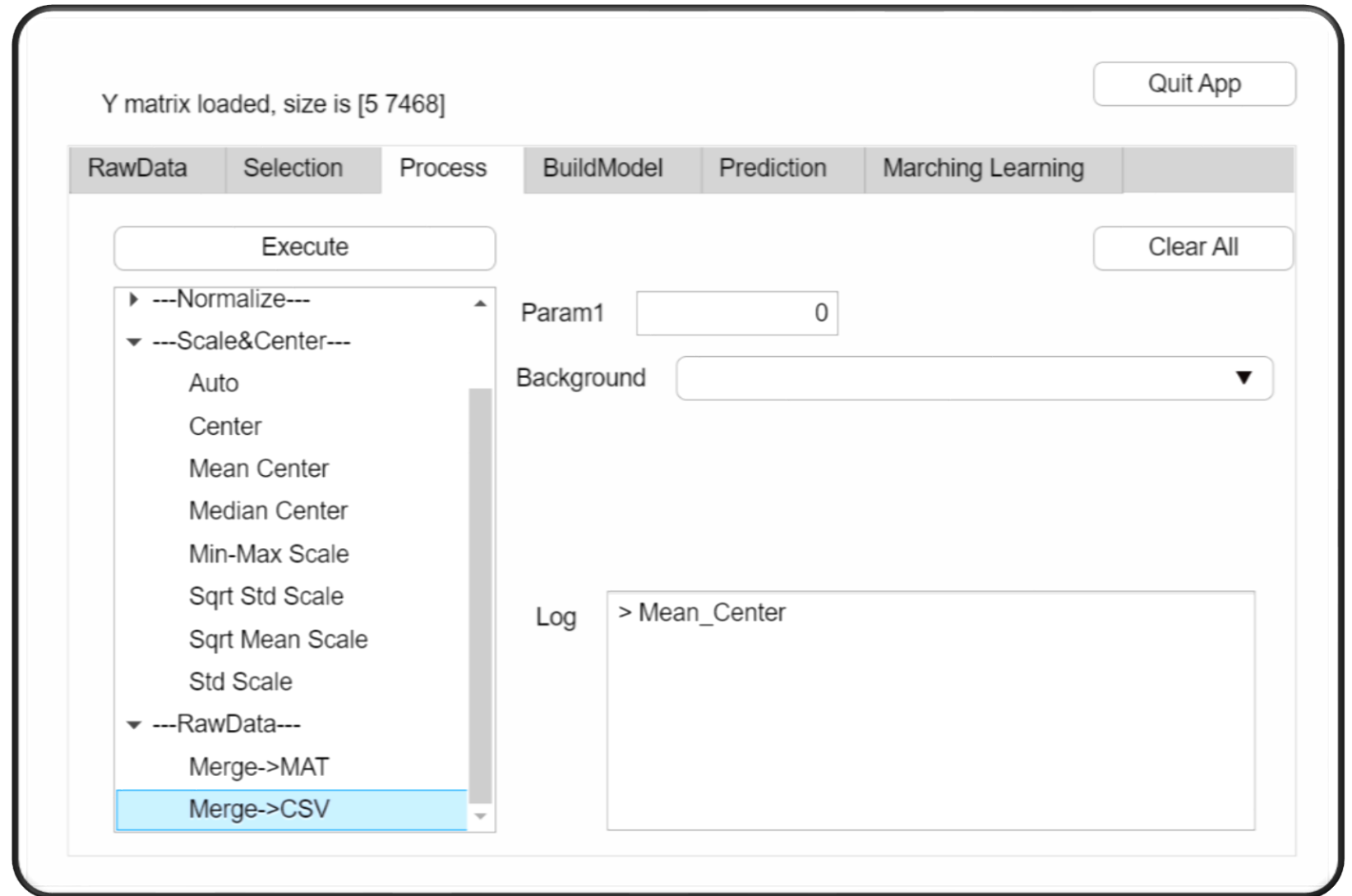
- PC list
- User sub-folder
- Keyword filter
- Date filter
- Data file selection
- Data preview
- Add to selection



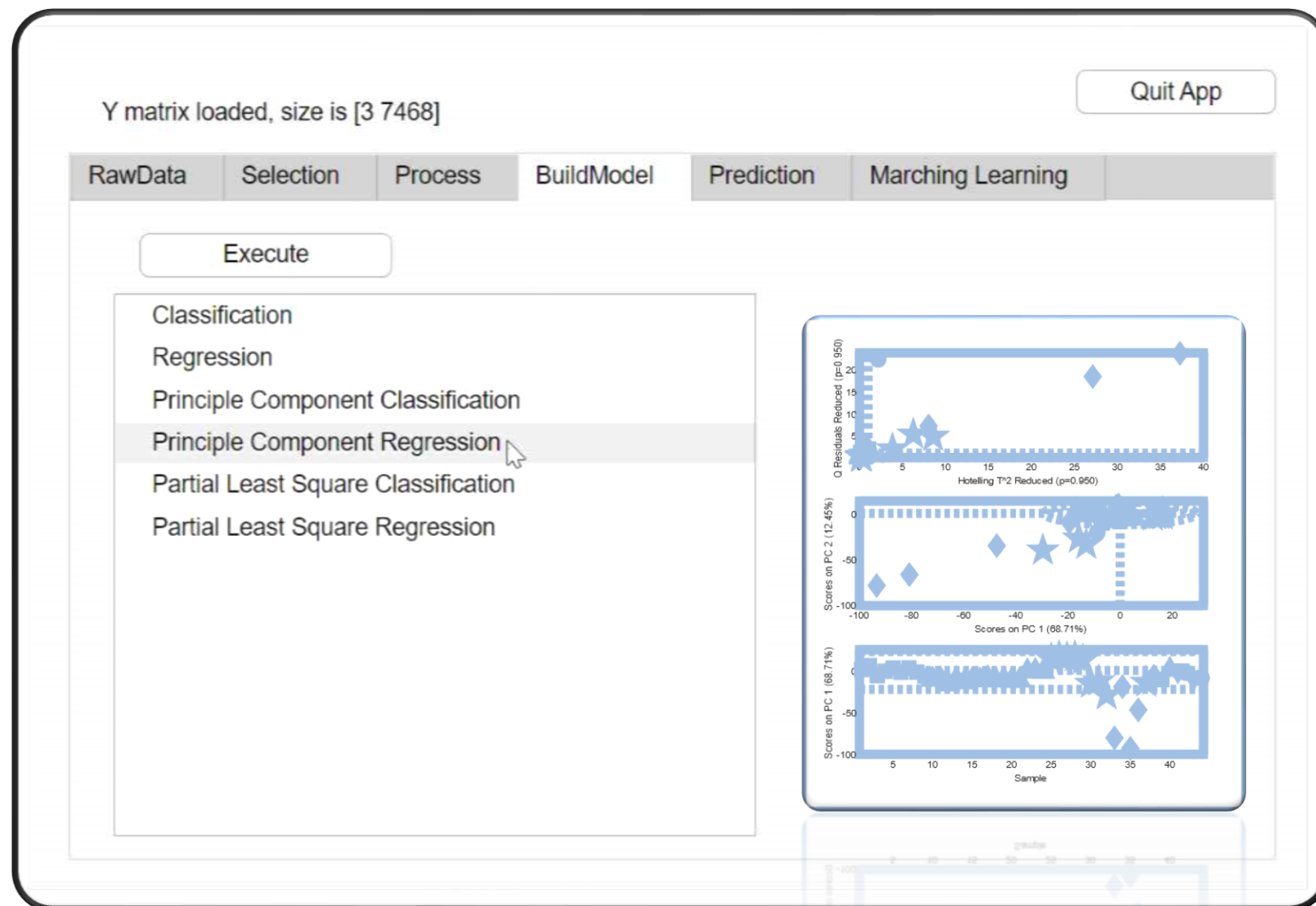
- Remove selection
- Check data size same or not
- Load Data to variable space



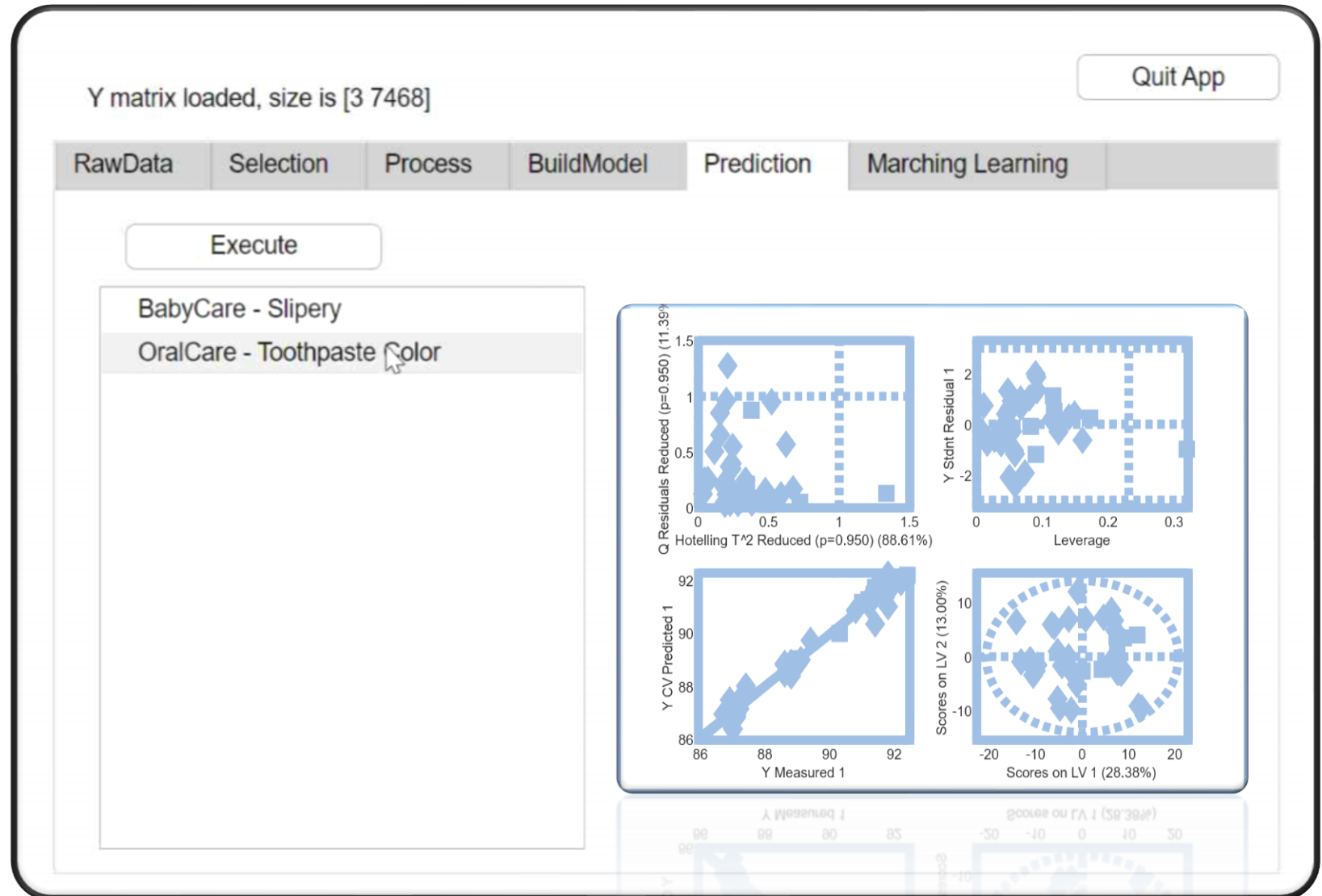
- Pre-Processing algorithms & parameter
- Log
- Clear all
- Data export



- Build models
- View output



- Load Data to predict



- Machine Learning common algorithms
- Basic model output
- Run several rounds with random parameters, to assess data readiness before data analyst involved**



Summary

- Web App Server
 - 7x24 service to customer
- App Designer
 - Graphic User Interface for easy operation
- Machine Learning
 - algorithms with randomized parameters to screen models
- Network
 - server access data via intranet
- User
 - access control, log, storage & reuse model, email etc.



Productive modeling platform for enterprise

MATLAB EXPO 2021

Thank you



© 2021 The MathWorks, Inc. MATLAB and Simulink are registered trademarks of The MathWorks, Inc. See [mathworks.com/trademarks](https://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.