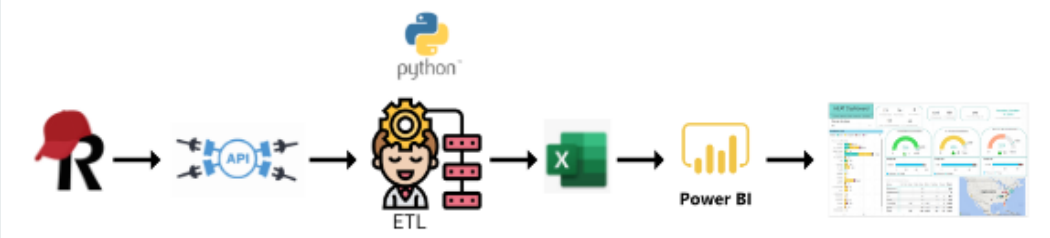
REDCap Interesting Use Case

Efficient Data Workflow Automation: Integrating REDCap API, Python, and PowerBI for Seamless & Real-time Multicenter Clinical Trials Reporting



Ally Qi, MPH
Clinical Trials Operations Manager
Vanderbilt Coordinating Center (VCC)
REDCap Fest
2/21/2024

Question:

How much time do you spend on creating or updating PPT slides per day/week/month?

Framing Eighteen coils in cerebral Aneurysms Trial (*FEAT*)

stryker

Multi-center, prospective, randomized trial comparing the utilization of *0.014-0.0155*" coils (**18 coils**) vs. smaller diameter coils (**standard 10 coils**) in mid-sized aneurysm treatment

651 participants enrelled across **25** sites in U.S.

Random **1:1** assignment

Primary Outcome: Angiographic Occlusion rate on follow-up angiography at 12-18 months

Secondary Outcomes

Treatment **related morbidity and mortality**, as measured by the NIH Stroke Scale

Packing density as measured by volumetric filling of the aneurysm

Modified Rankin Scale at 3-6- and 12-18-months post-coiling

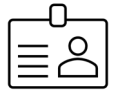
Re-hemorrhage and **re-treatment rates**

Challenges - What if?

Meetings rescheduled? 1x, 2x? Per week/month?

Cross-database data analysis

Special metrics of interest by PI



Site PI contact info

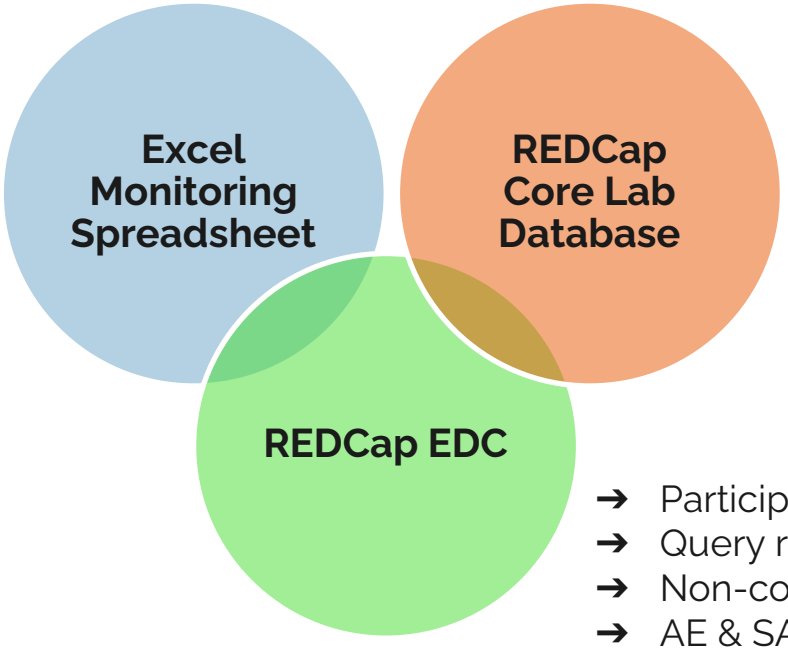
Accurate/up-to-date count of participant status

Unexpected questions to drill down into different categories or levels of data

Multi-site trial performance (PD & SAE & Query Resolution & Inventory)

Cross-database Data Wrangling

- ➔ Participant status
- ➔ Follow-up window
- ➔ Enrollment dates
- ➔ Monitoring status per instrument level
- ➔ Site contact info



- ➔ Core lab adjudication

- ➔ Participant data
- ➔ Query resolution
- ➔ Non-compliances
- ➔ AE & SAE

What're my options of data visual tools?

PowerPoint

PowerPoint slides

Power BI

tableau

PowerBI or Tableau

Applications

Project Dashboards

REDCap Dashboard

RStudio Shiny

R & R Shiny

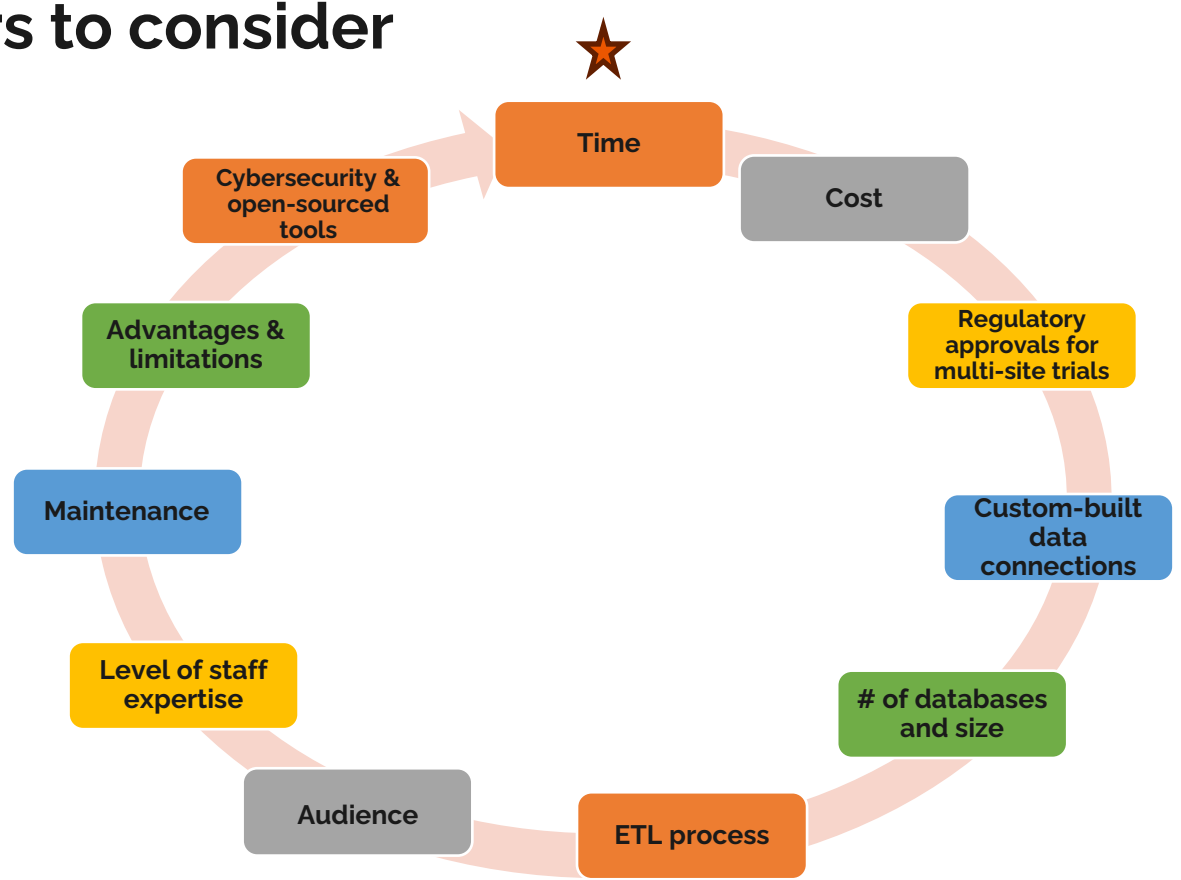
Google Data Studio

Google Data Studio

HTML CSS JavaScript

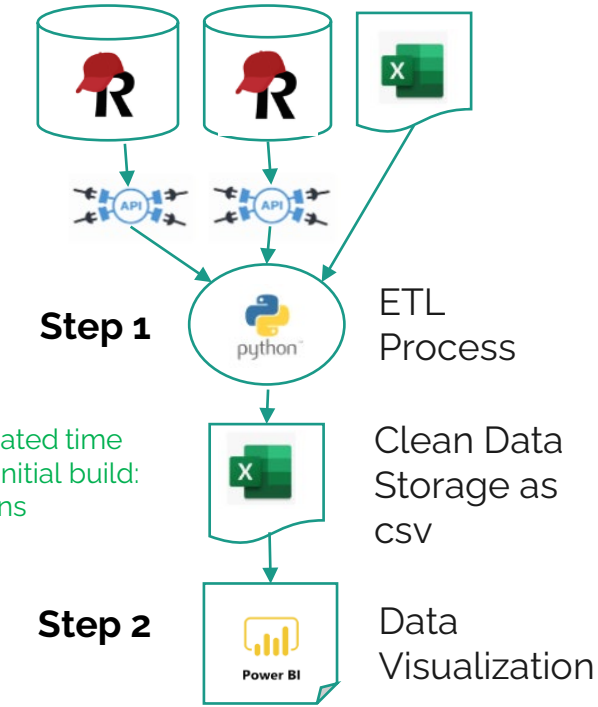
JavaScript & HTML & CSS

Factors to consider



FHIR

Solution! A Seamless Data Pipeline



Initial build:

1. In Python Jupyter Notebook, import data from 2 REDCap databases using API and 1 SharePoint Excel sheet to complete the ETL process. Then export clean data to flat csv files in a dedicated folder.
2. In PowerBI, build a dashboard to display key study metrics.

Benefits:

1. Free (limited version)*
2. Fast building time
3. Fastest loading time, closest to real time
4. Full automation
5. Reproducible & replicable
6. Secure
7. No regulatory approval needed
8. Low maintenance (debugging)

Key Metrics Identification

- **Data Safety Monitoring Plan (DSMP) and Study Protocol**
- Defines what can and can't be shared and protects blinding

- **Study Type**
- Investigational product(s): drugs or medical devices or interventions or others

- **Number of Time Points**
- Longitudinal vs single time point

- **Database structure**
- IP Arms (single vs multiple) and phases (start up, maintenance, and site closure)

- **Number of Sites**
- Single site vs Multi-sites

Key Metrics Identification

Enrollment:

at each time point, change over time

Participant Status

Incomplete: lost to follow-up, withdrawn, death, etc.

# of participants in window under each time point

Inclusion & Exclusion Criteria

Primary & Secondary Outcomes

Blinded vs open-labeled?

Safety Events:

AE & SAE

Non-compliances

REDCap Queries (open vs closed)

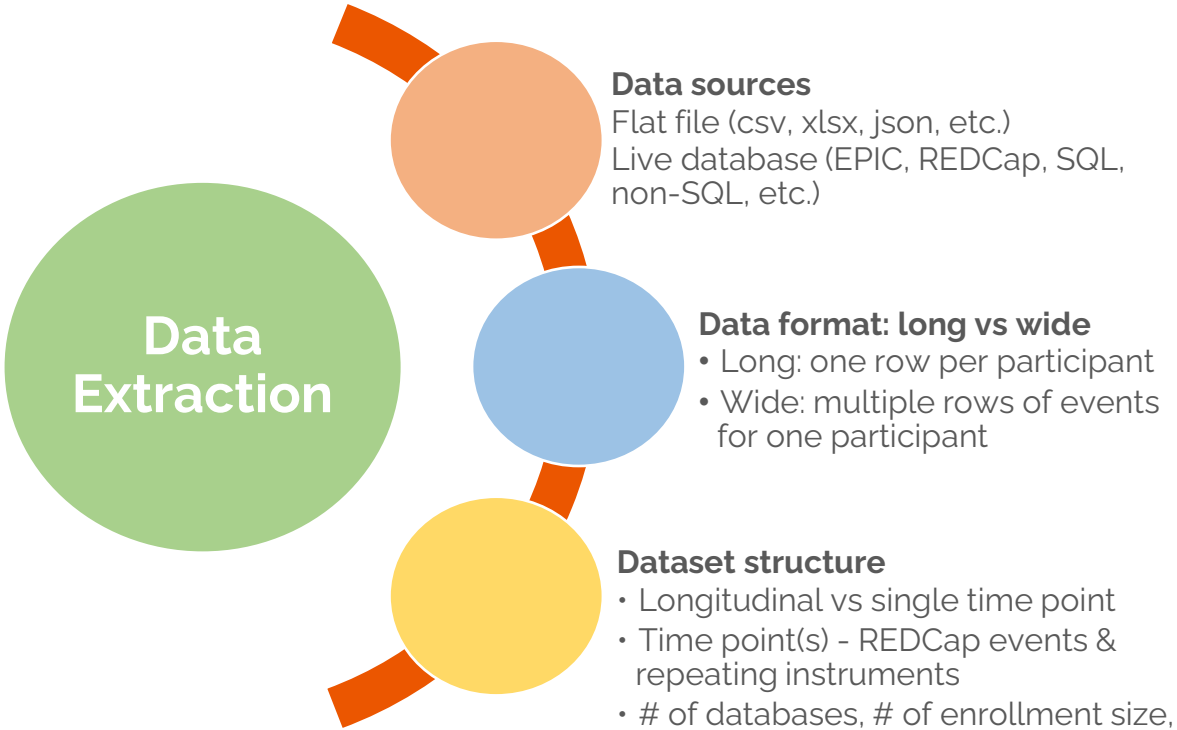
Median days to full resolution

Inventory Tracking: drugs, medical devices, imaging

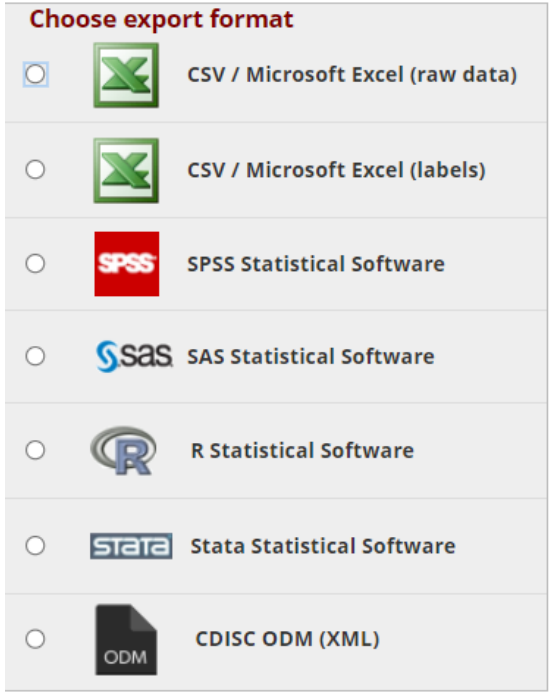
Site Activation Dates & Enrollment Speed

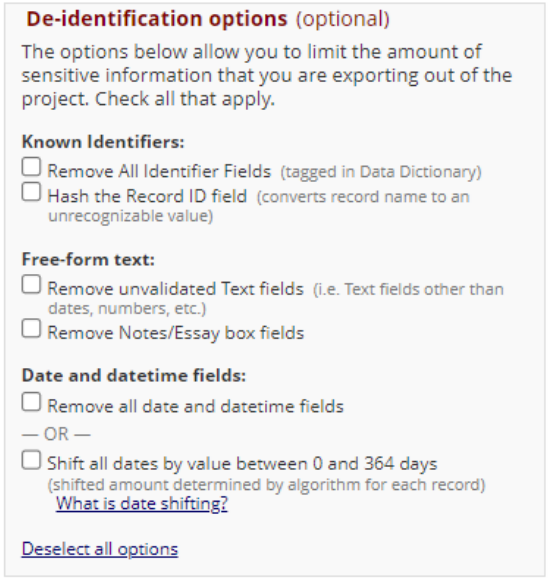
Special metrics of interest - PI

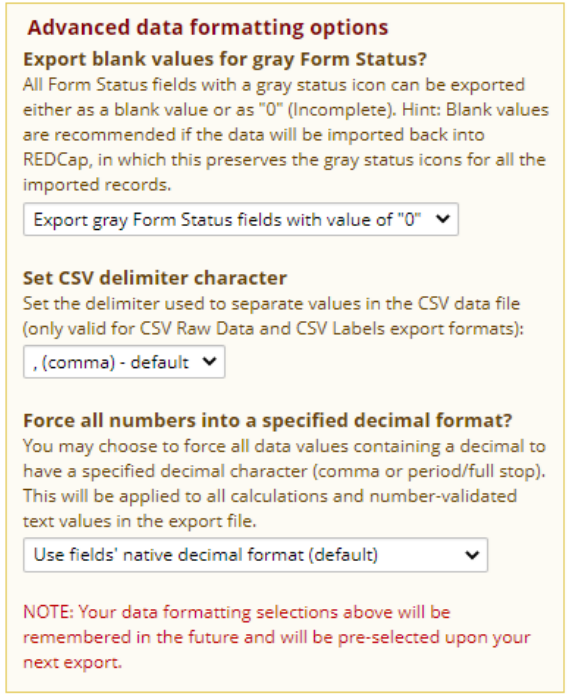
Data Extraction (ETL)



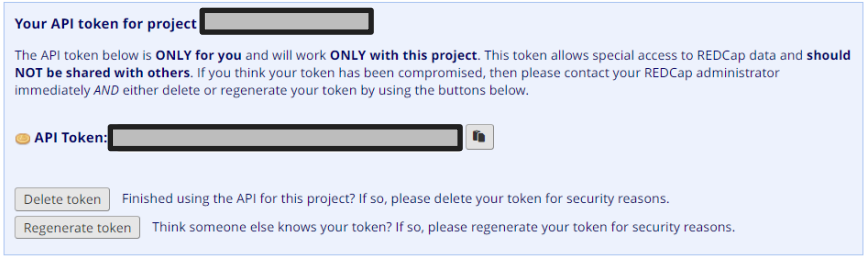
Basic REDCap Report Export







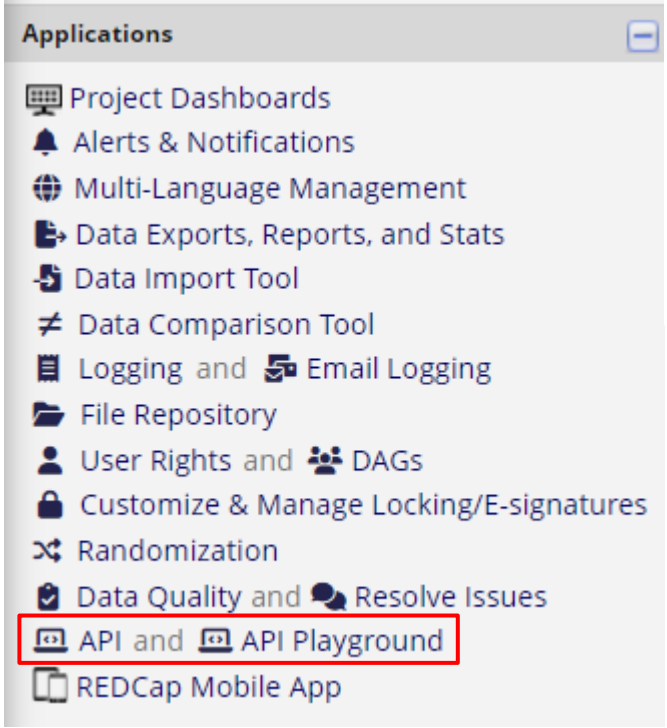
Intro to API (Application Programming Interface)

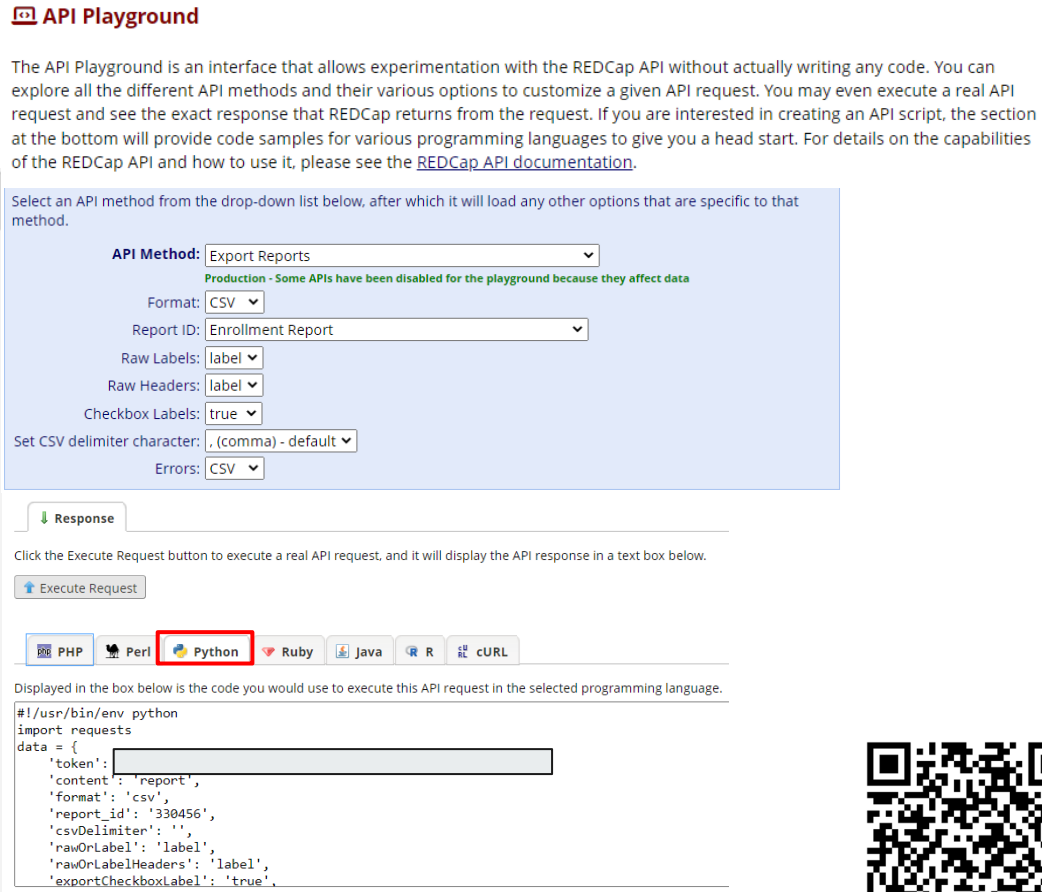


APIs (Application Programming Interfaces) are used to:

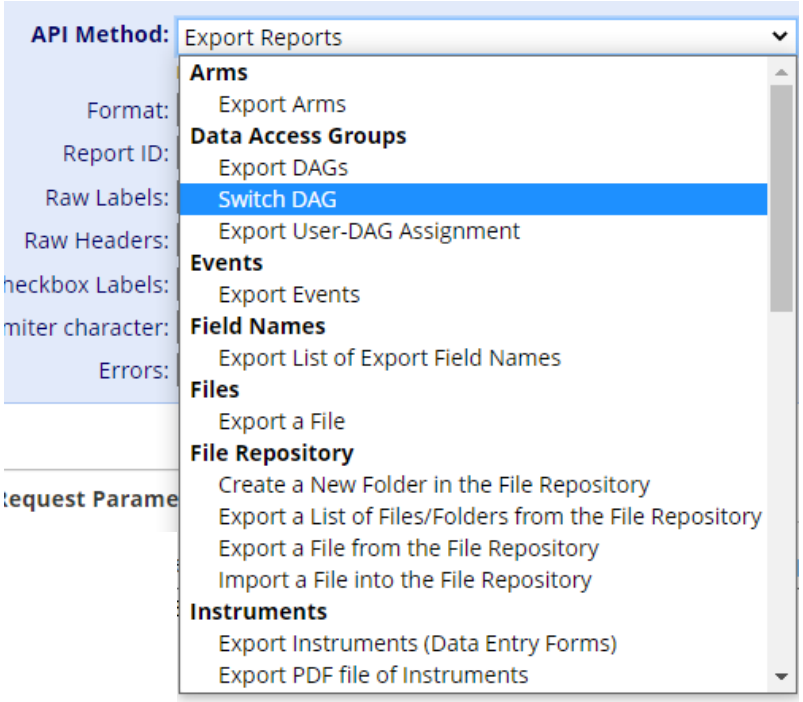
- provide a standardized way for data communication
- act as a bridge between different systems and allow them to exchange data seamlessly

REDCap API


Applications
Project Dashboards
Alerts & Notifications
Multi-Language Management
Data Exports, Reports, and Stats
Data Import Tool
Data Comparison Tool
Logging and Email Logging
File Repository
User Rights and DAGs
Customize & Manage Locking/E-signatures
Randomization
Data Quality and Resolve Issues
API and API Playground
REDCap Mobile App


API Playground
The API Playground is an interface that allows experimentation with the REDCap API without actually writing any code. You can explore all the different API methods and their various options to customize a given API request. You may even execute a real API request and see the exact response that REDCap returns from the request. If you are interested in creating an API script, the section at the bottom will provide code samples for various programming languages to give you a head start. For details on the capabilities of the REDCap API and how to use it, please see the REDCap API documentation.
Select an API method from the drop-down list below, after which it will load any other options that are specific to that method.
API Method: Export Reports
Production - Some APIs have been disabled for the playground because they affect data
Format: CSV
Report ID: Enrollment Report
Raw Labels: label
Raw Headers: label
Checkbox Labels: true
Set CSV delimiter character: , (comma) - default
Errors: CSV
Response
Click the Execute Request button to execute a real API request, and it will display the API response in a text box below.
Execute Request
PHP
Perl
Python
Ruby
Java
R
cURL
Displayed in the box below is the code you would use to execute this API request in the selected programming language.
#!/usr/bin/env python
import requests
data = {
'token':
'content': 'report',
'format': 'csv',
'report_id': '330456',
'csvDelimiter': '',
'rawOrLabel': 'label',
'rawOrLabelHeaders': 'label',
'exportCheckboxLabel': 'true',

What can I export from REDCap using API?



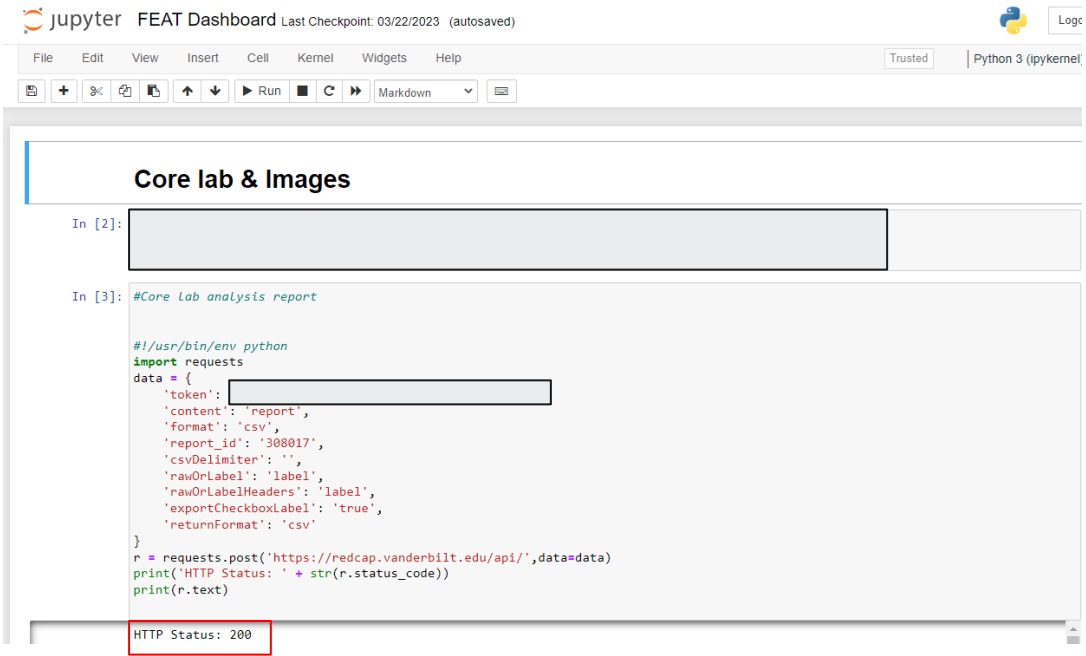
- Arms
- Data Access Groups
- Events & Field Names
- Files & File Repository
- Instruments
- Logging
- Metadata(data dictionary)
- Projects
- Records
- Repeating Instruments and Events
- Reports
- REDCap Version
- Survey
- Users & User Privileges & Roles

Intro to Python Pandas

Python Pandas:

- a powerful open-source data analysis and manipulation library
- provides easy-to-use data structures and data analysis tools
- provides a fast and efficient way to work with structured data
- has its intuitive and flexible API

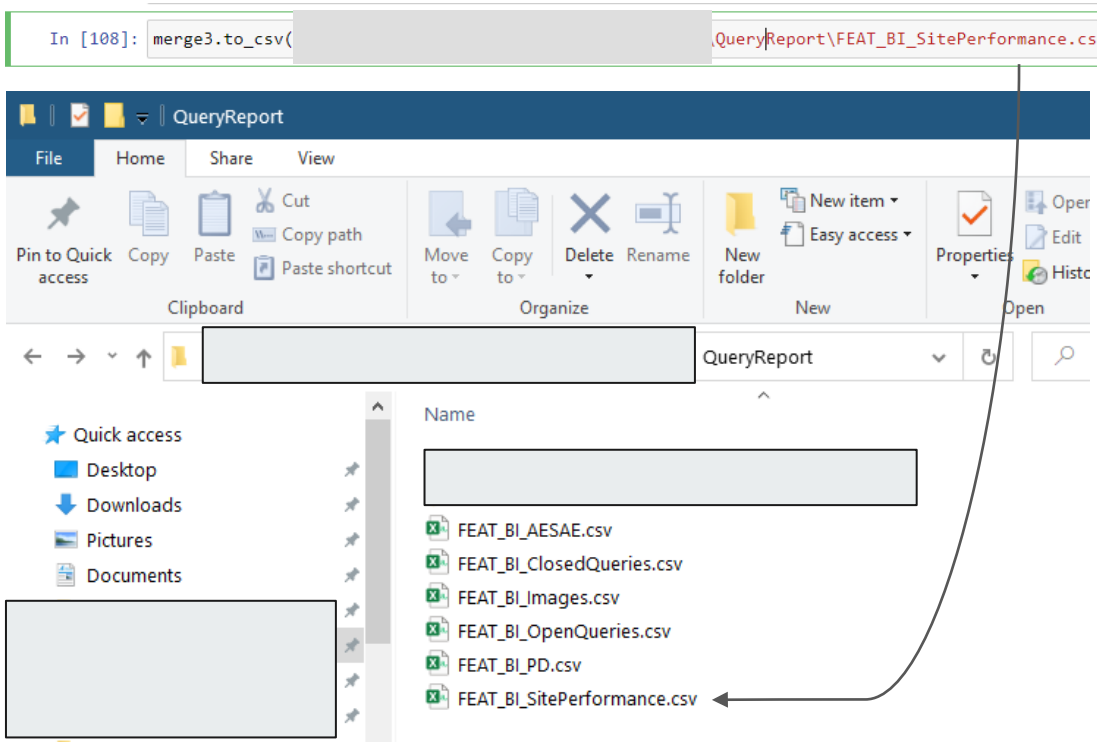
Data Transformation (ETL)



The API request was successful.

- ➔ 6 API Calls from 2 REDCap databases
- ➔ Each API call takes ~ 1 minute, depending on internet connection
- ➔ Import raw data and convert it to **Pandas dataframes** for data transformation
- ➔ Regular debugging is needed for maintenance

Data Load (ETL)



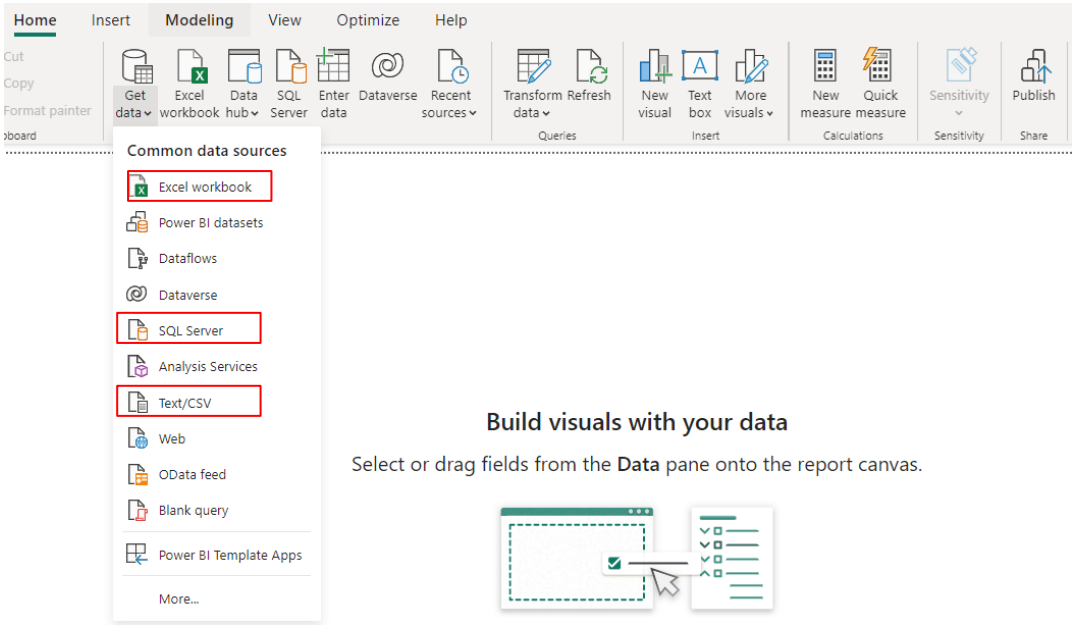
- Export clean data from Jupyter Notebook to csv files in a dedicated folder
- **No** timestamp in file names
- **Never** manually open/edit your clean/flat csv data files
  - Anytime if you catch a bug, you can correct the original data or Python code

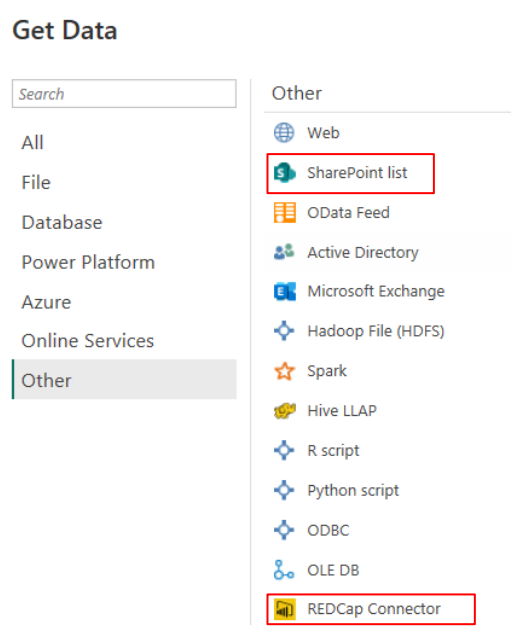
Intro to PowerBI Dashboard

PowerBI dashboard:

- a data visualization tool that allows users to create interactive and dynamic dashboards to analyze and present data
- connect to various data sources, such as Excel spreadsheets, databases, and cloud-based services, to collect and transform data into meaningful insights

PowerBI Dashboard Setup





Demo!

This demo dashboard used simulated datasets. All site info is publicly available on Clinicaltrials.gov.

Conclusion

How to connect all metrics together to tell a story?

- Updates: good news vs bad news
- Action items & decisions for the team
- Discussion topics

An iterative process until all needs and wants are met

- Feedback from audience

Less is more!

Time is of the essence!

What's the true value of this?

- Using **free* and existing** technologies to build a **seamless, fast and automatic** data pipeline that provides **accurate and real-time** data for team members (esp. PIs) to make **timely** decisions.

Questions? Feedback? More?

Contact Me

Email: ally.qi@vumc.org

Office: 615-875-0619

Thank you!

VCC family:

Stephanie Smith, Jessica Marlin, David McKeel, Kelly Walsh, John Graves, Courtney Jordan, and LSS.

Leadership:

Amanda Bistran-Hall, Krista Vermillion, Jess Collins, Bree Burks, Dr. Michael Froehler

For more info about VCC:
https://vcc.vumc.org

VANDERBILT COORDINATING CENTER

Mount Sinai

National PI: Dr. J Mocco, MD, MS
Kalmon D. Post Professor and Senior Vice Chair
Cerebrovascular Center Director
Department of Neurological Surgery
Mount Sinai Health System