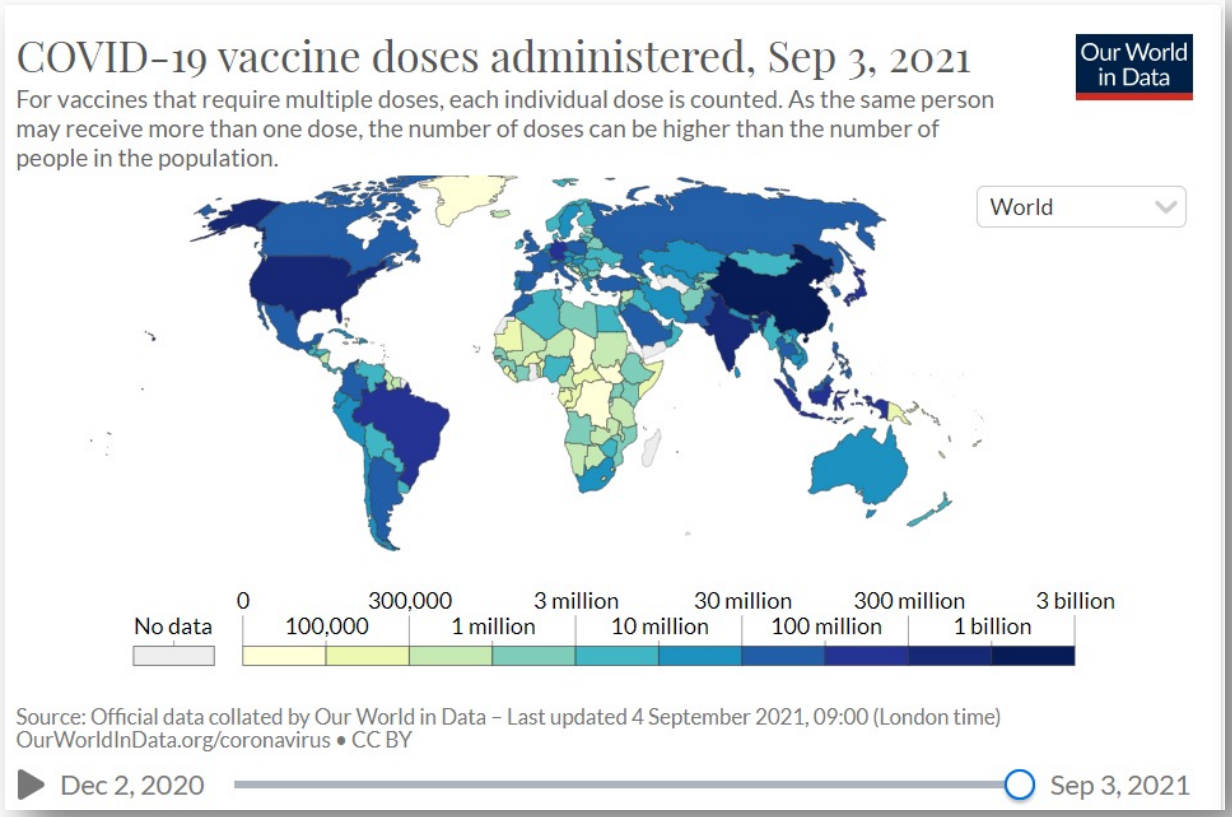
Visualizations in Clinical Trials using Power BI and Python

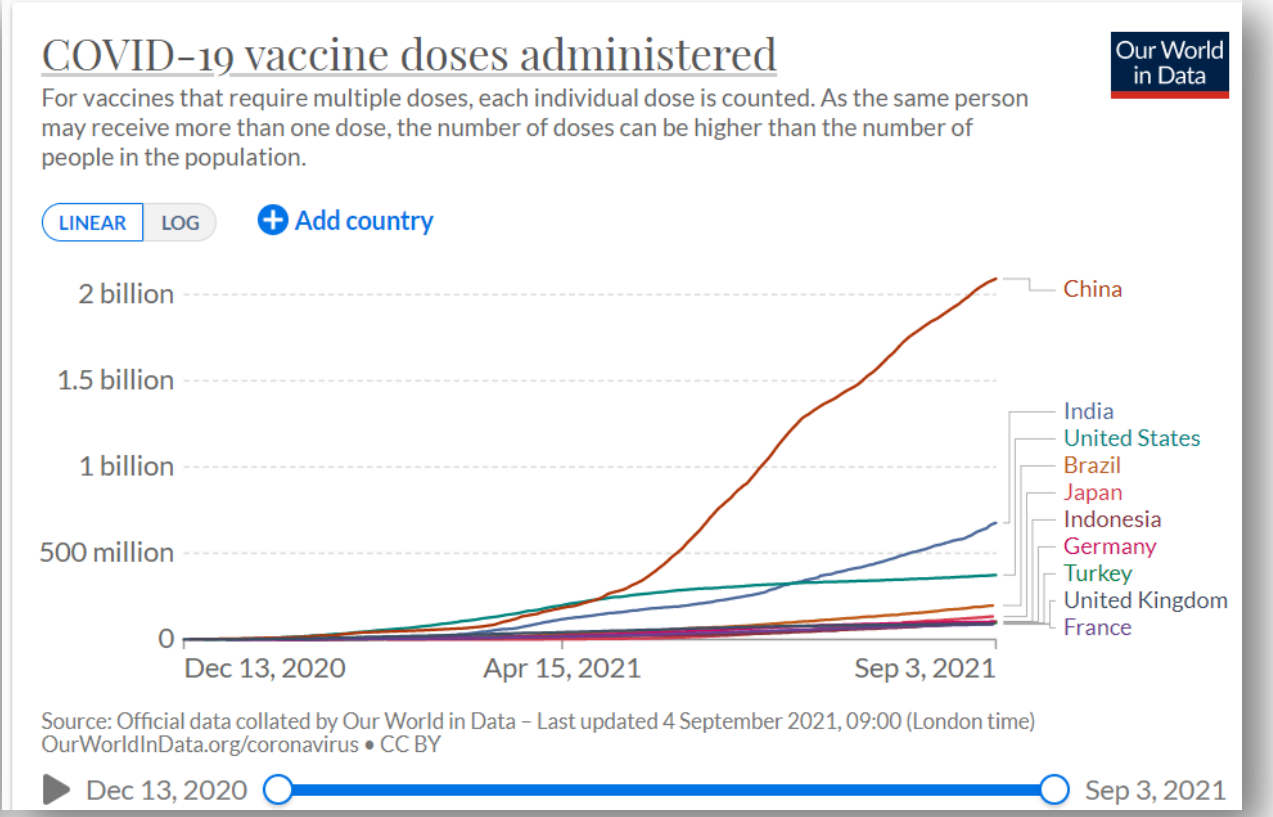
Bhavin Solanki,
Computational Scientist, Oncology Computational Science, Bayer
India Autumn SDE - September 18, 2021

Disclaimer

The information and opinions expressed in this presentation are solely my own and do not express the views or opinions of my employer, Bayer.

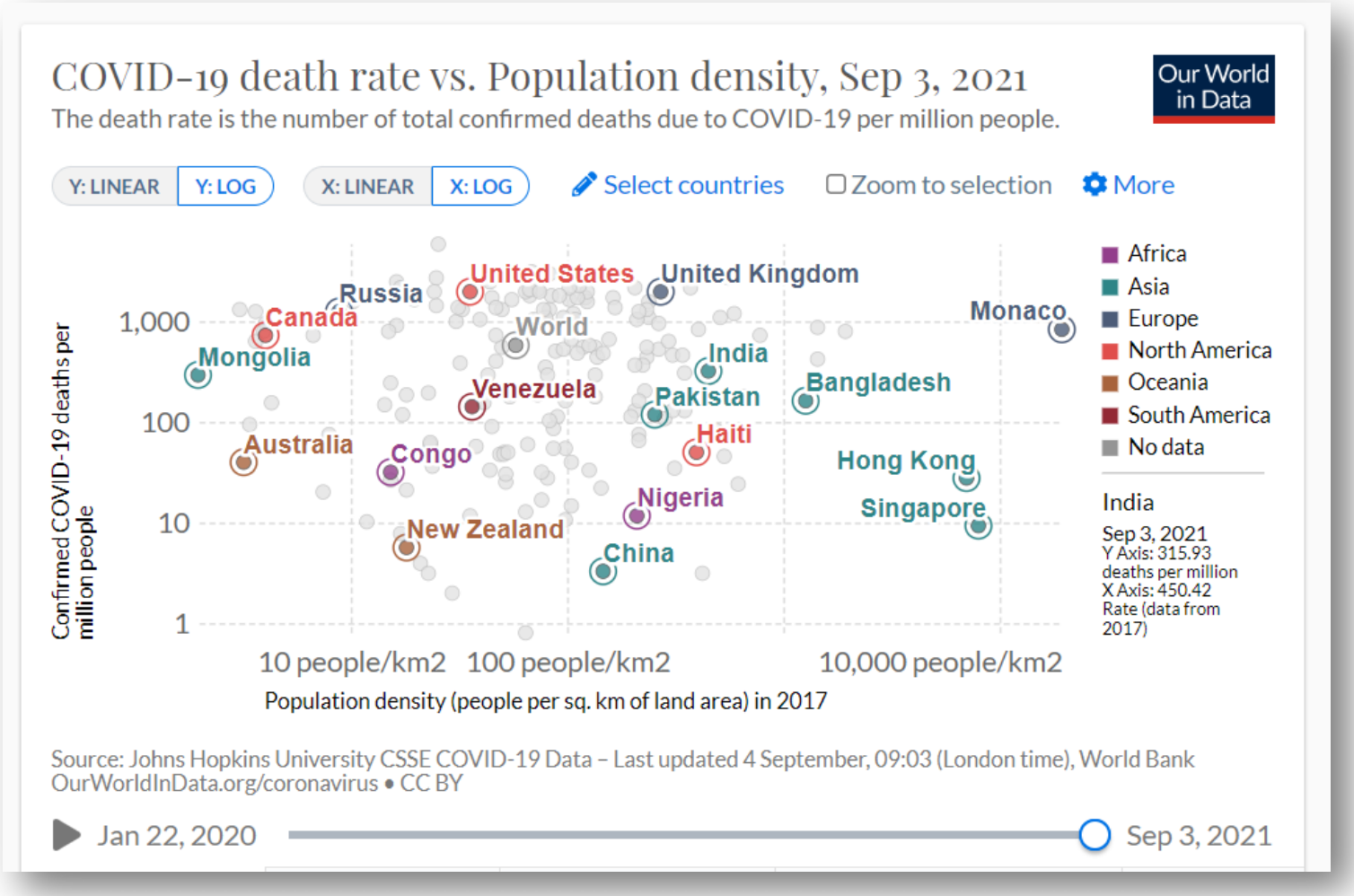
Let’s start with some visualizations





Why visualizations?

- Quick and easy way of understanding data
- Most efficient holistic way of looking at data
- Enable us to draw quick insights and faster decision making
- Makes the data more natural for the human mind to comprehend
- Makes it easier to identify trends, patterns, and outliers
- Communicates information in a fast and effective way
- Sometimes answers the questions that we don't know to ask
- Promotes storytelling



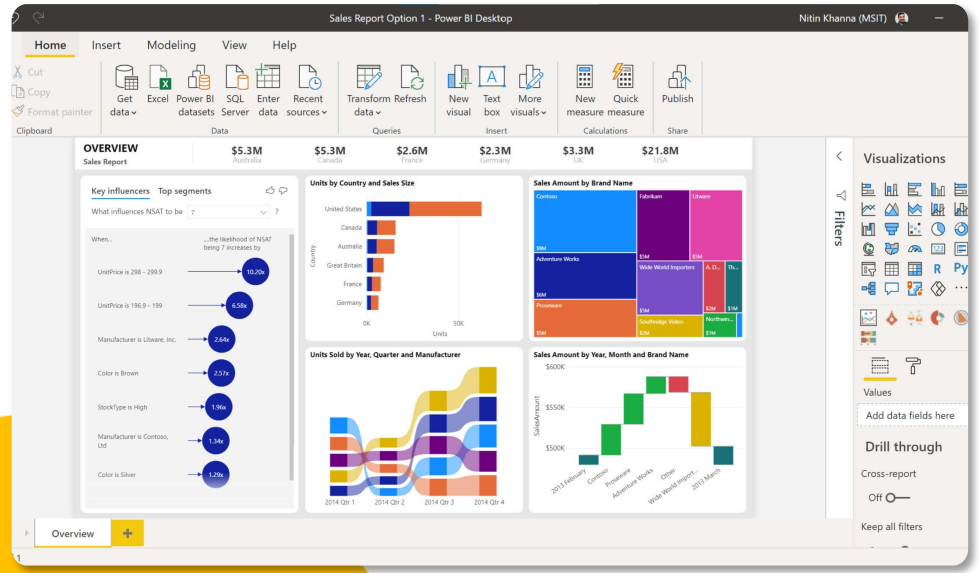
Popular Data Visualization Tools in Clinical Research Domain

No-code/low-code

Programming Languages

Microsoft

Power BI



Pros

- Free to use, easy to learn
- No code/Low code solution
- Interactive Visualizations
- Intuitive
- Less Error Prone
- Distribution is easy (Requires license)
- Easy data wrangling
- Carries data
- Seamless integration with Microsoft platforms

Cons (Vs Python)

- Limitations in customization
- Handling very large datasets can be challenging
- Easy to learn, complex to master

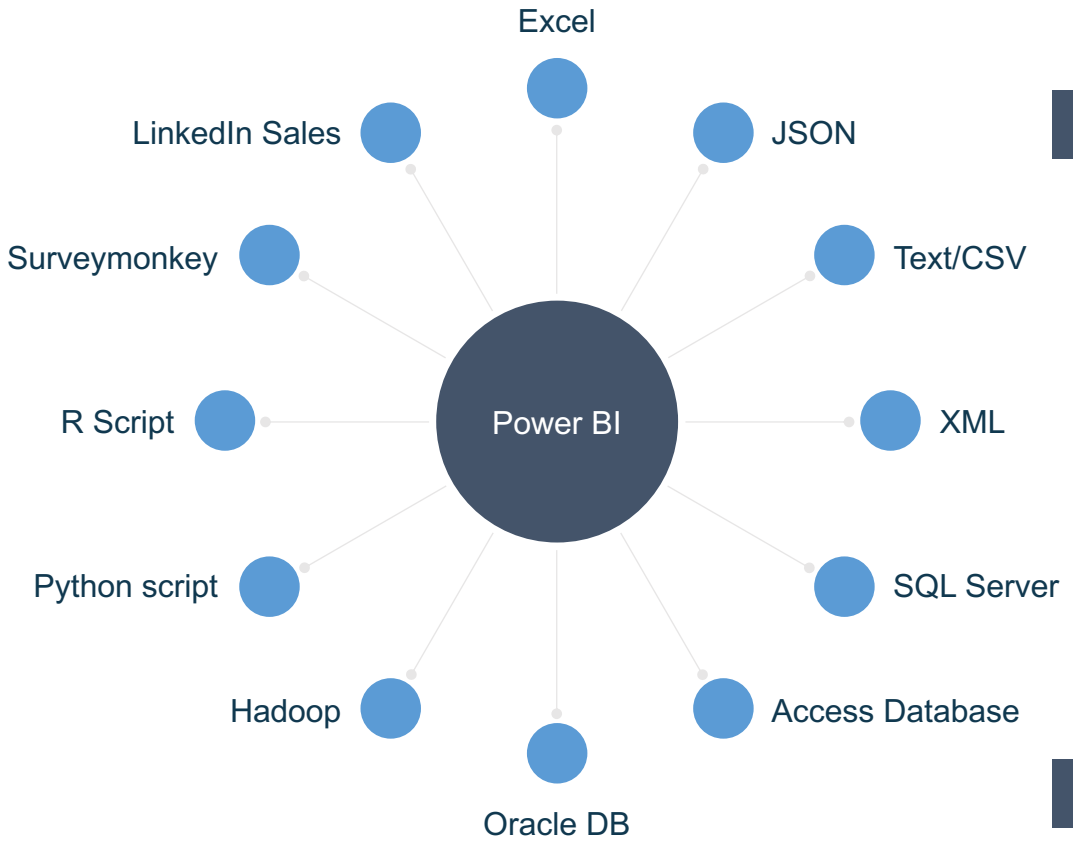
Resources

- Download from this link

Data Connections

Online services

Files



Almost anything

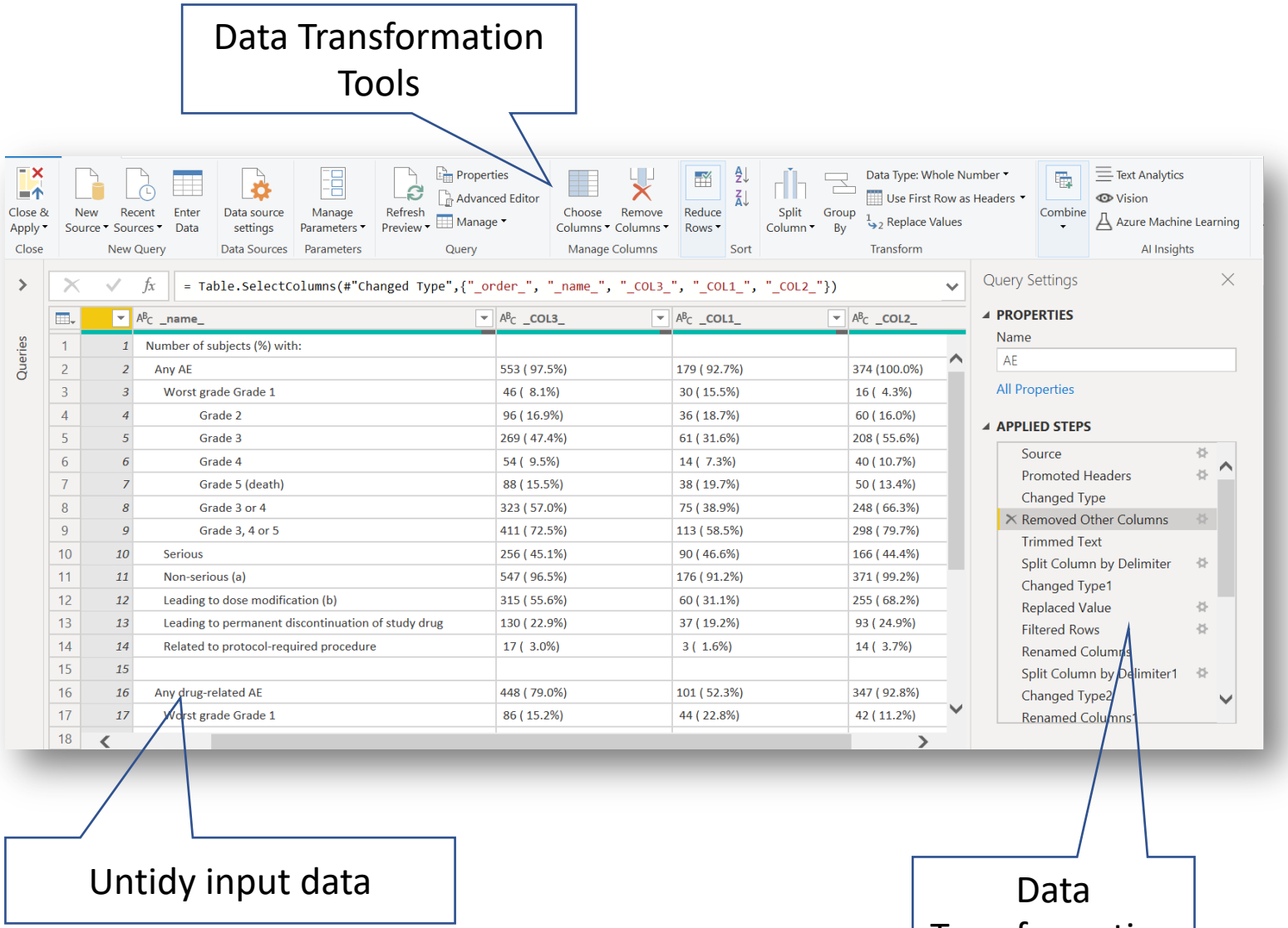
Databases

Use case

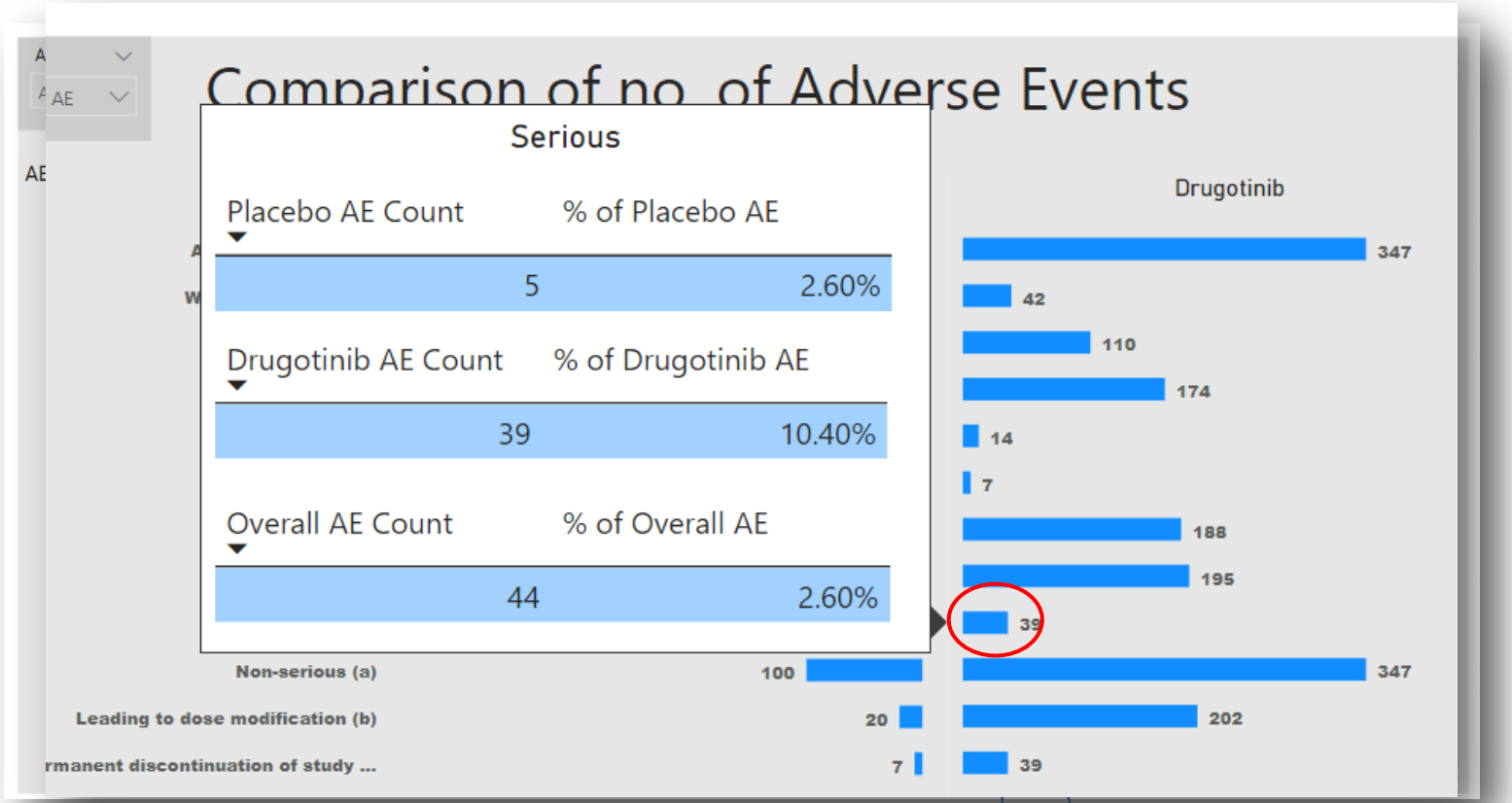
From **Adverse events summary table** to **visualization** in less than 30 minutes with Power BI

- Convert the SAS dataset to .csv file and import data in Power BI
- Transform data
- Create visualizations

Power BI Data Transformation



Adverse Events Visualization



Hover mouse pointer to see more details

What else can Power BI do?

- Drill through the visualizations
- Automated data refresh
- Send automated mails
- Accessible through smartphone app
- Embed the dashboard in apps, SharePoint, Teams
- Export the dashboard report in ppt or pdf
- Custom Visualizations
- Power BI analytics (Automated) and narratives
- Export data in excel

Learn from

Challenges, projects, prizes, datasets

Questions?

Pros

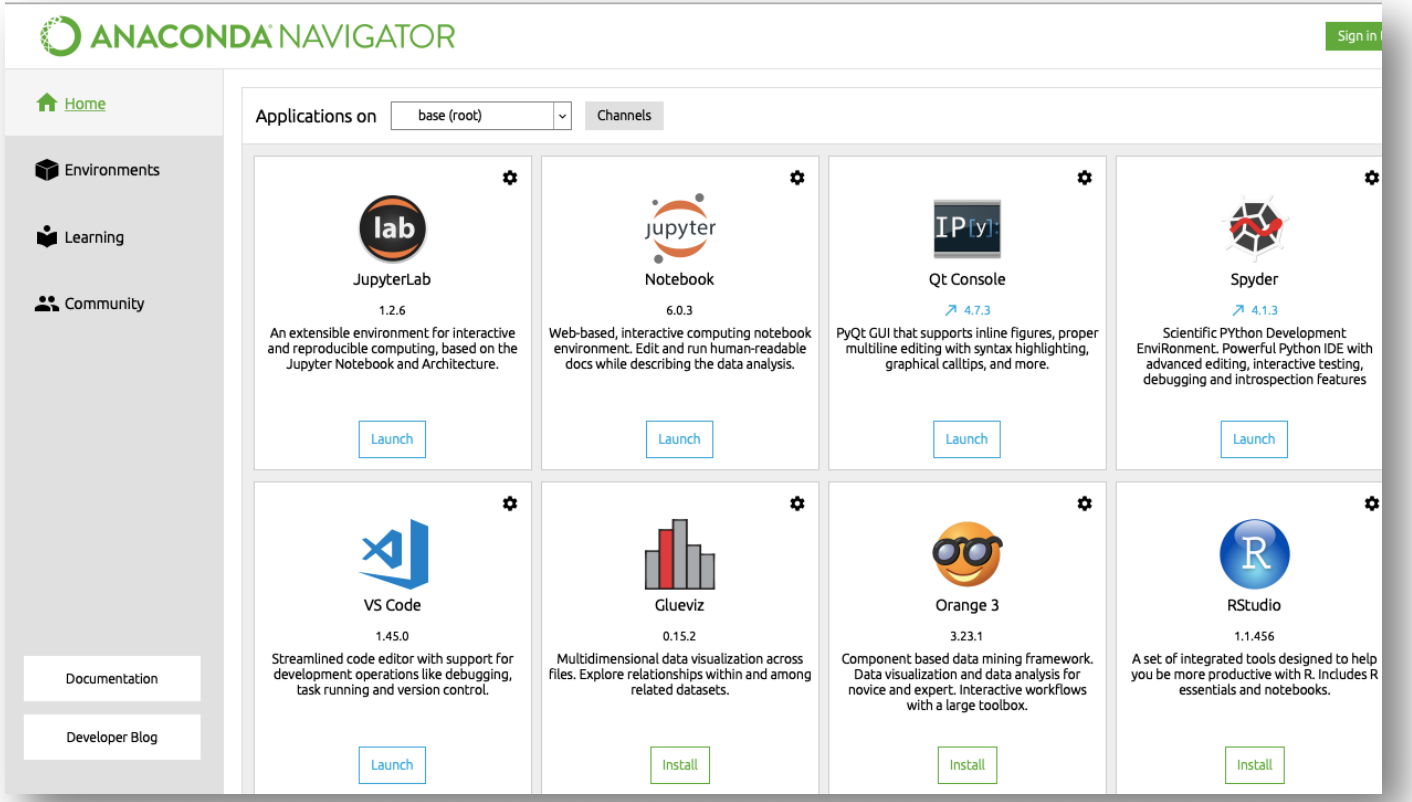
- “The” tool for data science
- Free and open source
- Simple and easy to use
- Static as well as Interactive visualizations
- It’s a complete programming language
- Scalability
- Various libraries/packages
- Latest machine learning libraries

Cons (Vs Power BI)

- Quite a lot of coding
- More Error Prone
- Distribution and deployment is not easy
- Data is stored separately

Resources

- Download from this link



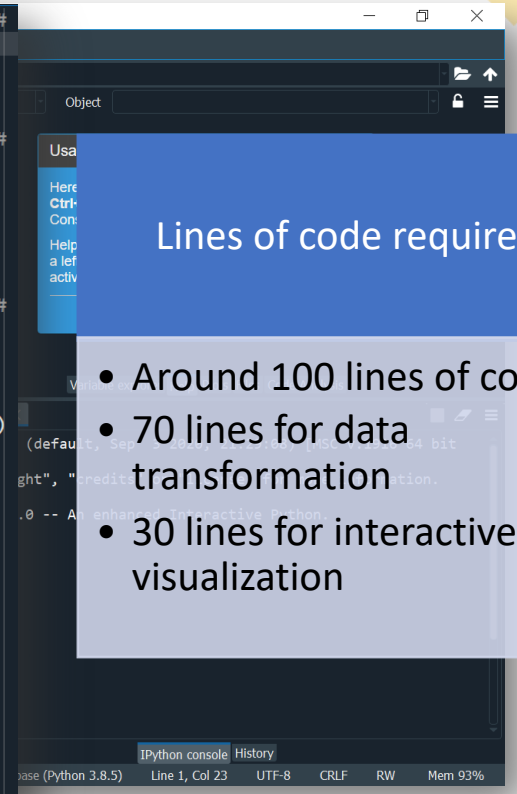
Use case

Adverse Events Visualization in python in an hour

- Read data from SAS dataset
- Transform the data to make them tidy
- Use *“plotly”* library to create interactive visualizations
- Use “*matplotlib*” or “*seaborn*” library to create static visualizations

Coding in python using Spyder

```
141  ##############################################################################
142
143  #https://plotly.com/python/subplots/
144
145  #Interactive visualizations
146  ##############################################################################
147
148  import os
149  import plotly.offline as pyo
150  import plotly.graph_objs as go
151  from plotly.subplots import make_subplots
152
153  ##############################################################################
154
155  from TopLineResults_Preprocessing import df_anyAE
156
157  Placebo = df_anyAE['Placebo_1']
158  Placebo_size = df_anyAE['Placebo_2'].apply(lambda x: 10 if x <= 10 else 10+x/2)
159  AE = df_anyAE['AE']
160
161  fig = make_subplots(rows=1,
162                      cols=2,
163                      shared_yaxes=True,
164                      subplot_titles=("Any AE",
165                                      "Any drug-related AE"))
166  #Adding traces
167  fig.add_trace(go.Scatter(
168      x=Placebo,
169      y=AE,
170      marker=dict(color="crimson", size=Placebo_size),
171      mode="markers",
172      name="Placebo"
173  ),row=1, col=1)
174
```



Lines of code required

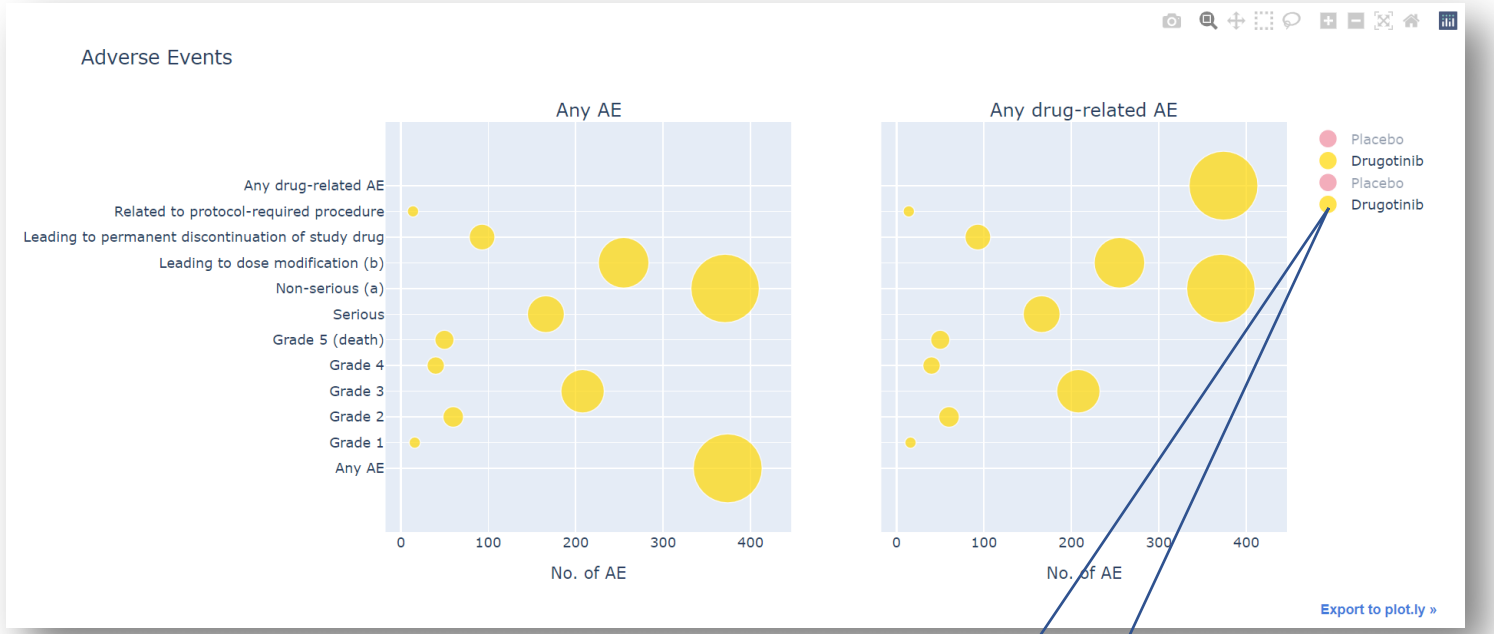
- Around 100 lines of code
- 70 lines for data transformation
- 30 lines for interactive visualization

Write data transformation code in editor

Adverse Events Visualization in .html file

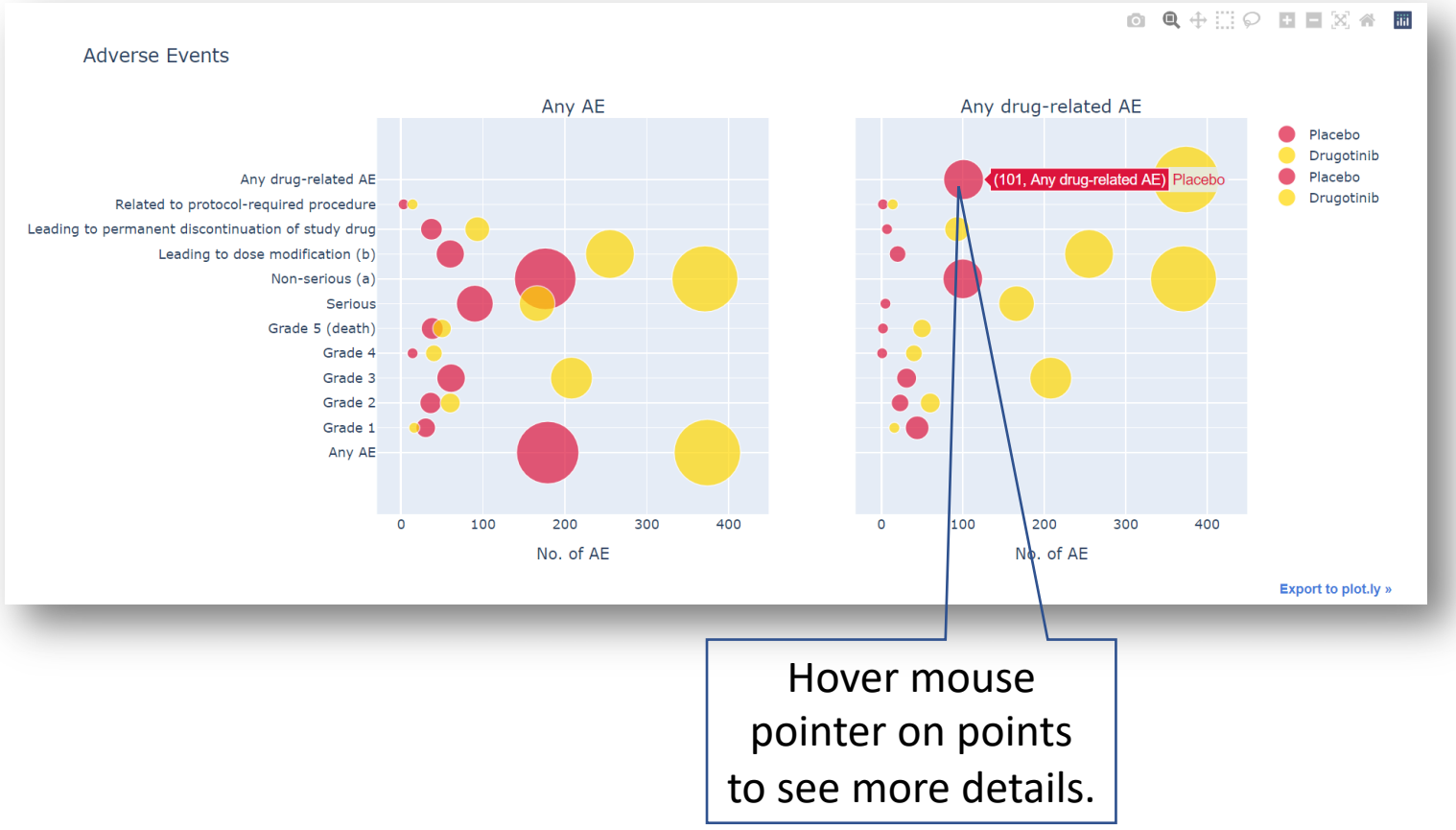
Export static image, zoom in and zoom out, focus on specific area etc.

Interactivity of plotly visualizations



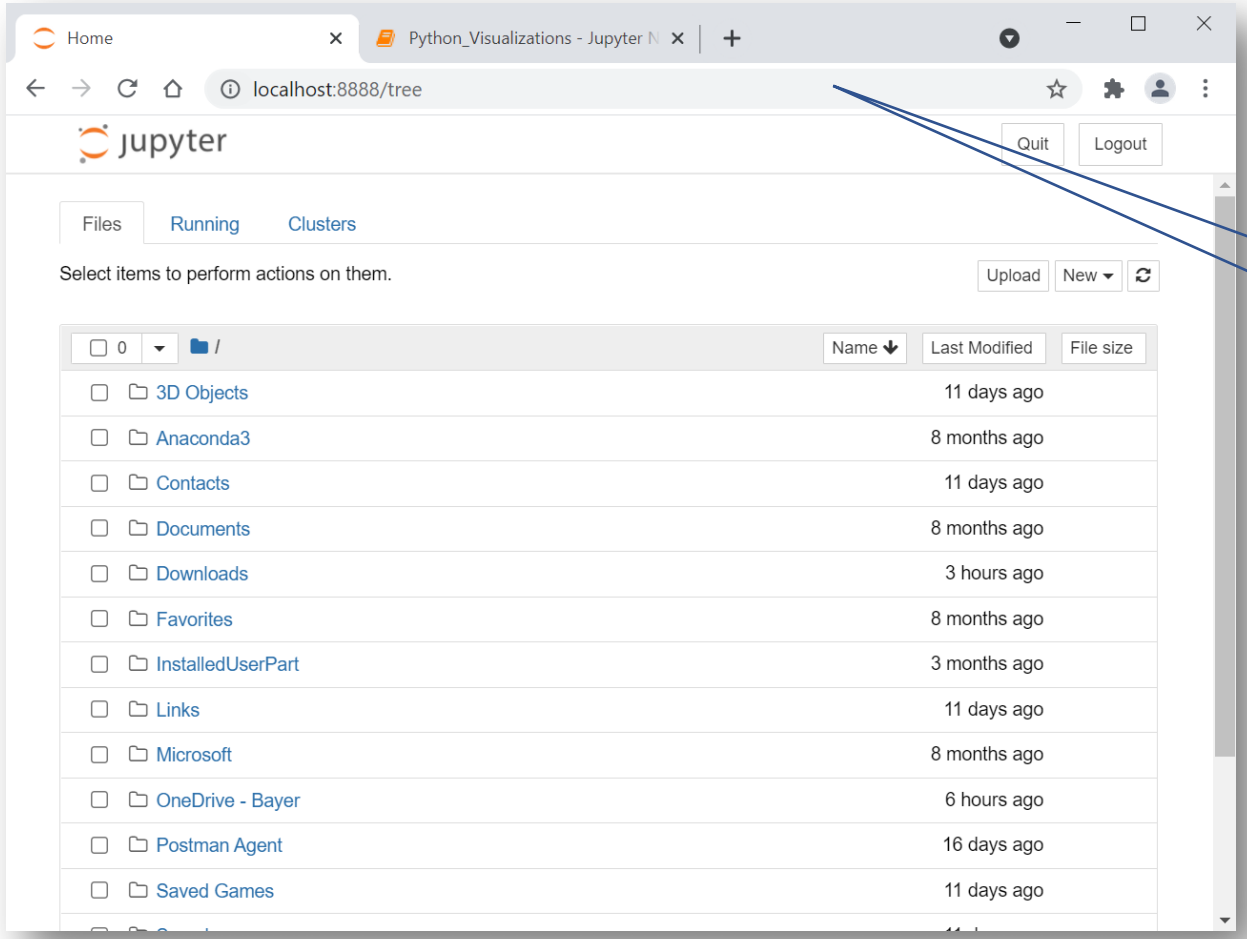
Select one arm to see only the chart for that arm

Tooltips in plotly visualizations



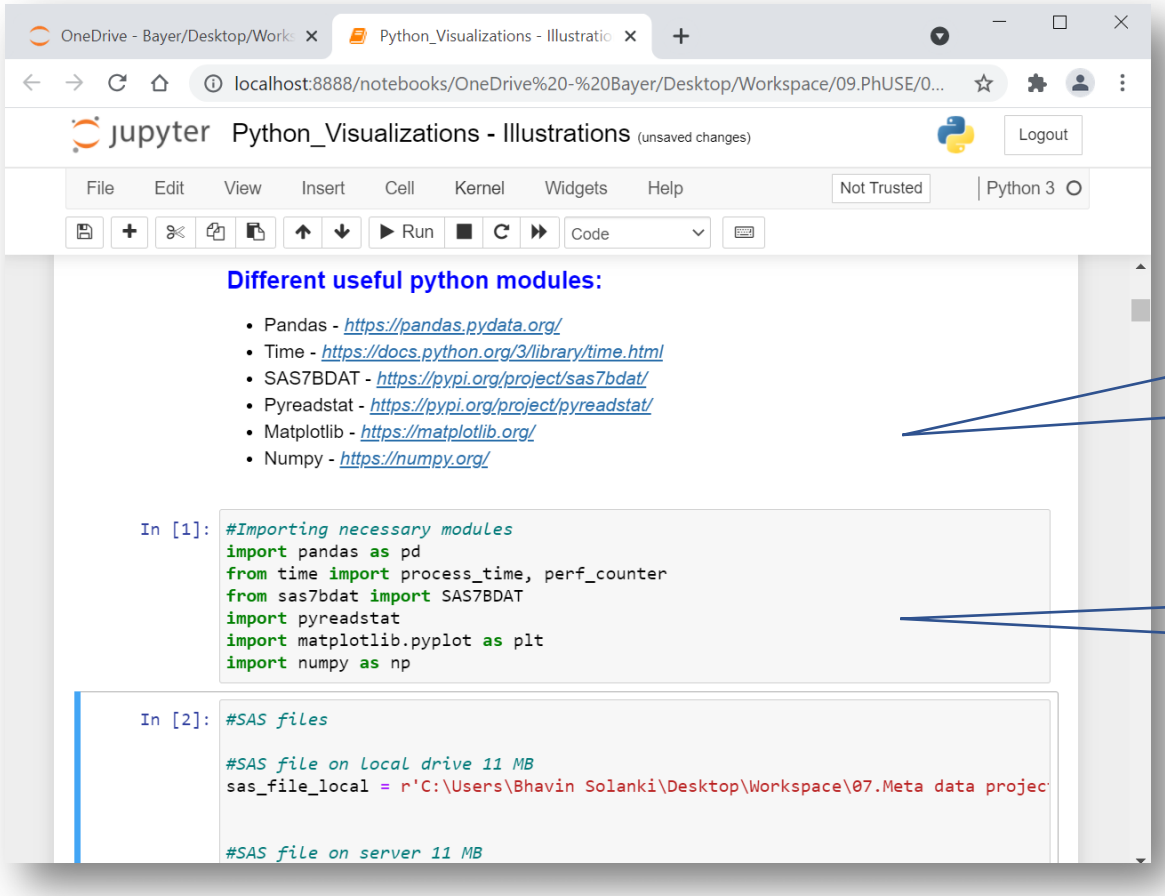
Hover mouse pointer on points to see more details.

Visualization in jupyter Notebook



Opens in a browser

Code in Jupyter notebook



Equivalent to R Markdown file

Write code in cells

Output in Jupyter notebook

```
In [16]: #Cleaning and making the SAS dataset tidy

         #Adverse events
         ae_sas_file = r'C:\Users\Bhavin Solanki\Desktop\Workspace\01.Topline Results\03.SAS Datasets\t_ae.sas7bdat'
         df = pd.read_sas(ae_sas_file, encoding="latin1")

         #checking the dataframe
         df
```

Out[16]:

	order	_name_	_COL3_	_COL1_	_COL2_	fline	error	_orderN	_obsid_	_1_b1	_1_b2	_1_b3
0	1.0	Number of subjects (%) with:	NaN	NaN	NaN	1.0	0.0	1.0	1.0	1.0	1.0	1.0
1	2.0	Any AE	553 (97.5%)	179 (92.7%)	374 (100.0%)	2.0	0.0	2.0	2.0	1.0	1.0	1.0
2	3.0	Worst grade Grade 1	46 (8.1%)	30 (15.5%)	16 (4.3%)	3.0	0.0	3.0	3.0	1.0	1.0	1.0
3	4.0	Grade 2	96 (16.9%)	36 (18.7%)	60 (16.0%)	4.0	0.0	4.0	4.0	1.0	1.0	1.0
4	5.0	Grade 3	269 (47.4%)	61 (31.6%)	208 (55.6%)	5.0	0.0	5.0	5.0	1.0	1.0	1.0
5	6.0	Grade 4	54 (9.5%)	14 (7.3%)	40 (10.7%)	6.0	0.0	6.0	6.0	1.0	1.0	1.0
6	7.0	Grade 5 (death)	88 (15.5%)	38 (19.7%)	50 (13.4%)	7.0	0.0	7.0	7.0	1.0	1.0	1.0
7	8.0	Grade 3 or 4	323 (57.0%)	75 (38.9%)	248 (66.3%)	8.0	0.0	8.0	8.0	1.0	1.0	1.0
8	9.0	Grade 3, 4 or 5	411 (72.5%)	113 (58.5%)	298 (79.7%)	9.0	0.0	9.0	9.0	1.0	1.0	1.0
9	10.0	Serious	256 (45.1%)	90 (46.6%)	166 (44.4%)	10.0	0.0	10.0	10.0	1.0	1.0	1.0

Prints the result just below the “code” cell

Output in Jupyter notebook

In [17]:
```
#List of treatments in the clinical study
trt = ['Placebo','Drugotinib']

#Columns to keep
cols_to_keep = ['_name_','_COL1_','_COL2_']

df1 = df[cols_to_keep].dropna() #Method chaining
df1['_name_'] = df1._name_.str.strip()
df1.columns = ["AE"] + trt

#Checking first few lines
df1.head()
```

Out[17]:

	AE	Placebo	Drugotinib
1	Any AE	179 (92.7%)	374 (100.0%)
2	Worst grade Grade 1	30 (15.5%)	16 (4.3%)
3	Grade 2	36 (18.7%)	60 (16.0%)
4	Grade 3	61 (31.6%)	208 (55.6%)
5	Grade 4	14 (7.3%)	40 (10.7%)

Quick summary table

Static Visualization in Jupyter notebook

jupyter Python_Visualizations - Illustrations (autosaved)

Logout

File Edit View Insert Cell Kernel Widgets Help Not Trusted Python 3

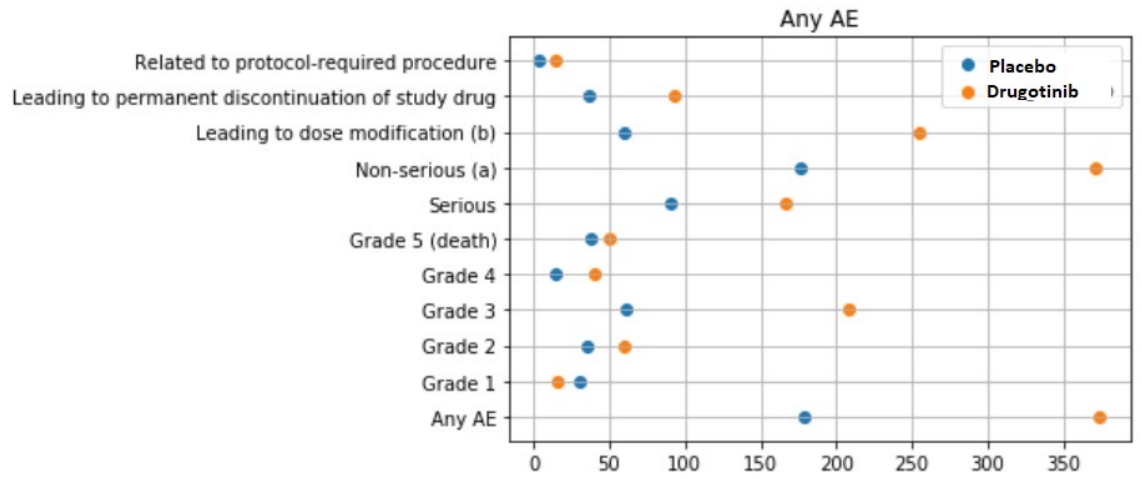
Run Code

```
plt.title('Any AE')

#Switching on legends
plt.legend()

#Displaying the grid for easy comparison
plt.grid()

#Display the figure
plt.show()
```



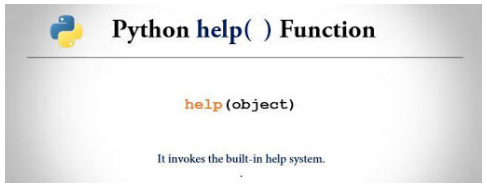
What else can python do?

- Data Science
  - Handling Data
  - Visualizations
  - Machine learning
  - Scrap websites
- Web applications
- Desktop applications
- Websites
- Mobile applications
- Just about anything that excludes low-level machine functions and writing operating systems

Learn from

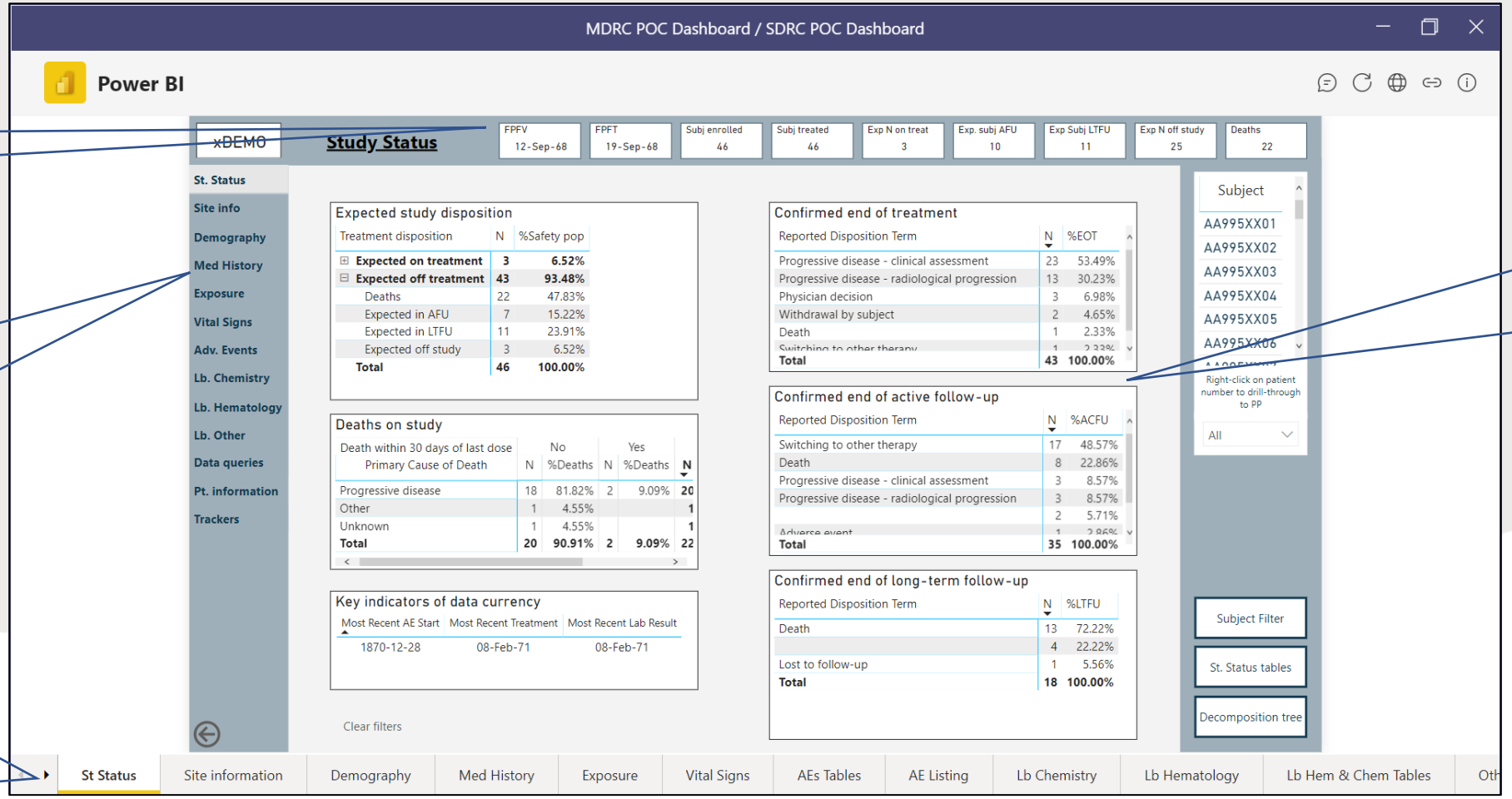
Challenges, projects, prizes, datasets

Questions?



Use case
Dashboard created in Power BI for Clinical Trial Scientists for Medical Review

Power BI Dashboard embedded in Teams Channel



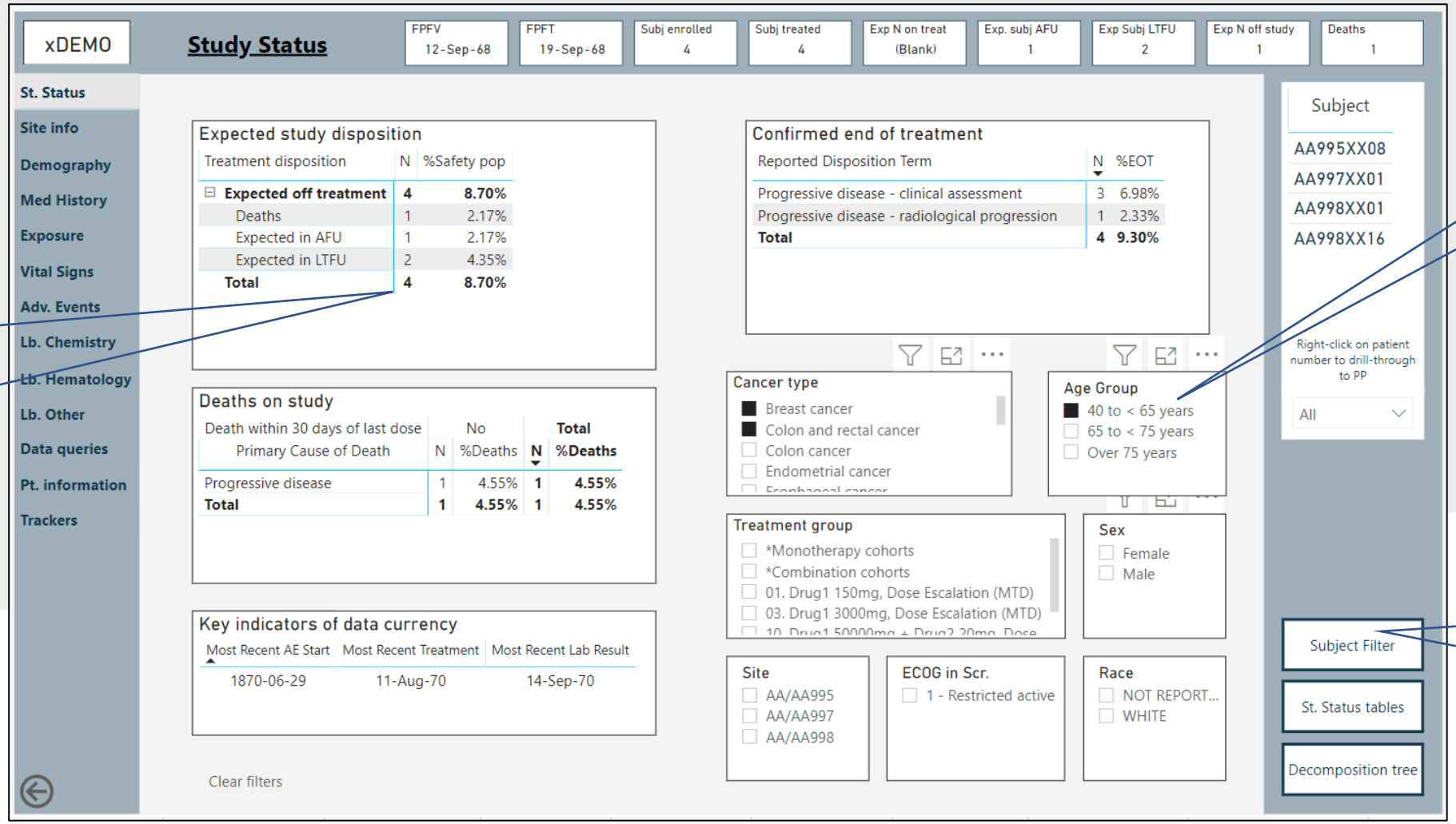
Information Cards

Navigation links created for easy navigation

Different report pages

Interactive and expandable Table visualizations

Filtering data using slicers

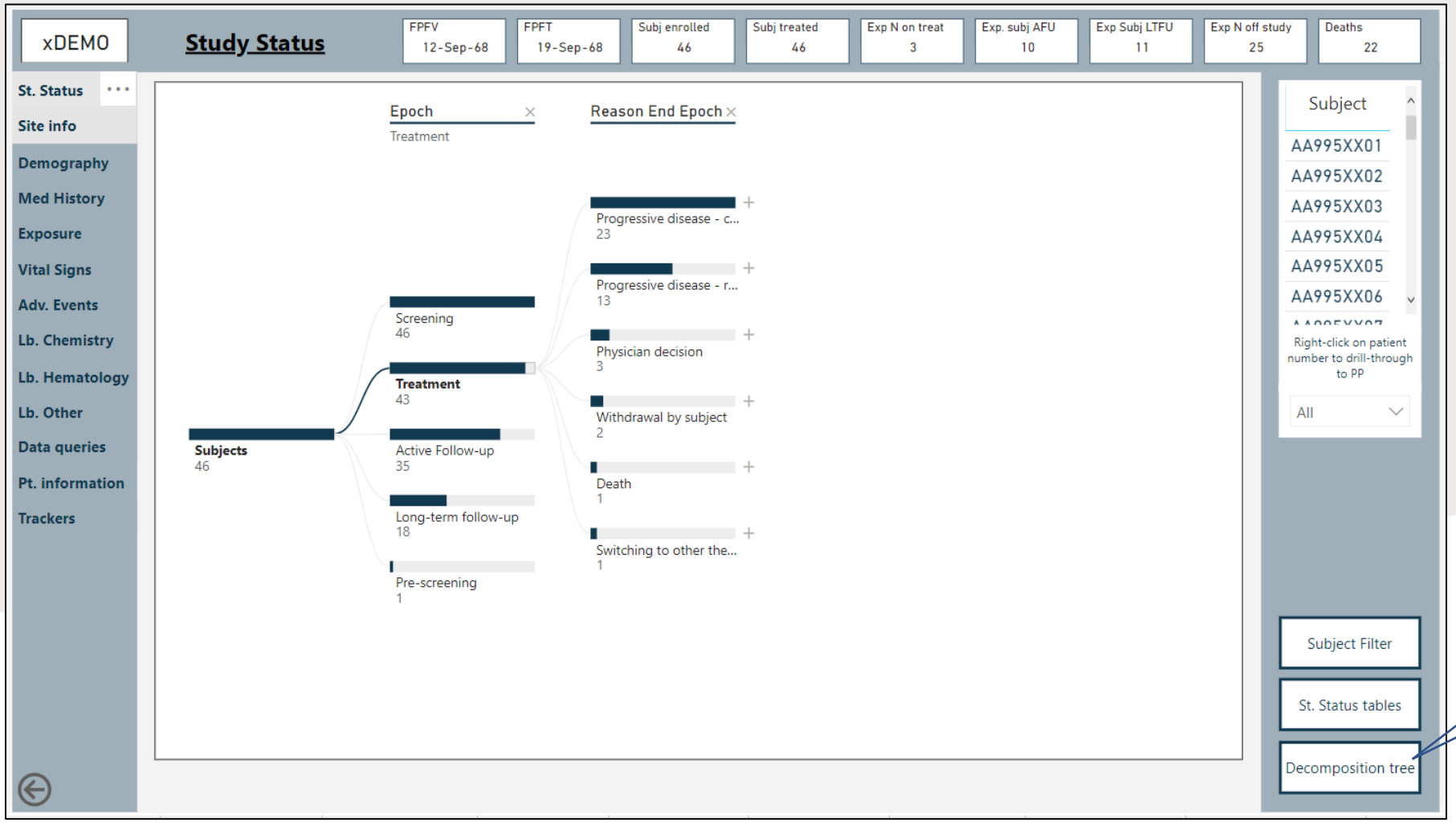


All the data tables and cards are automatically updated

Selected Filters

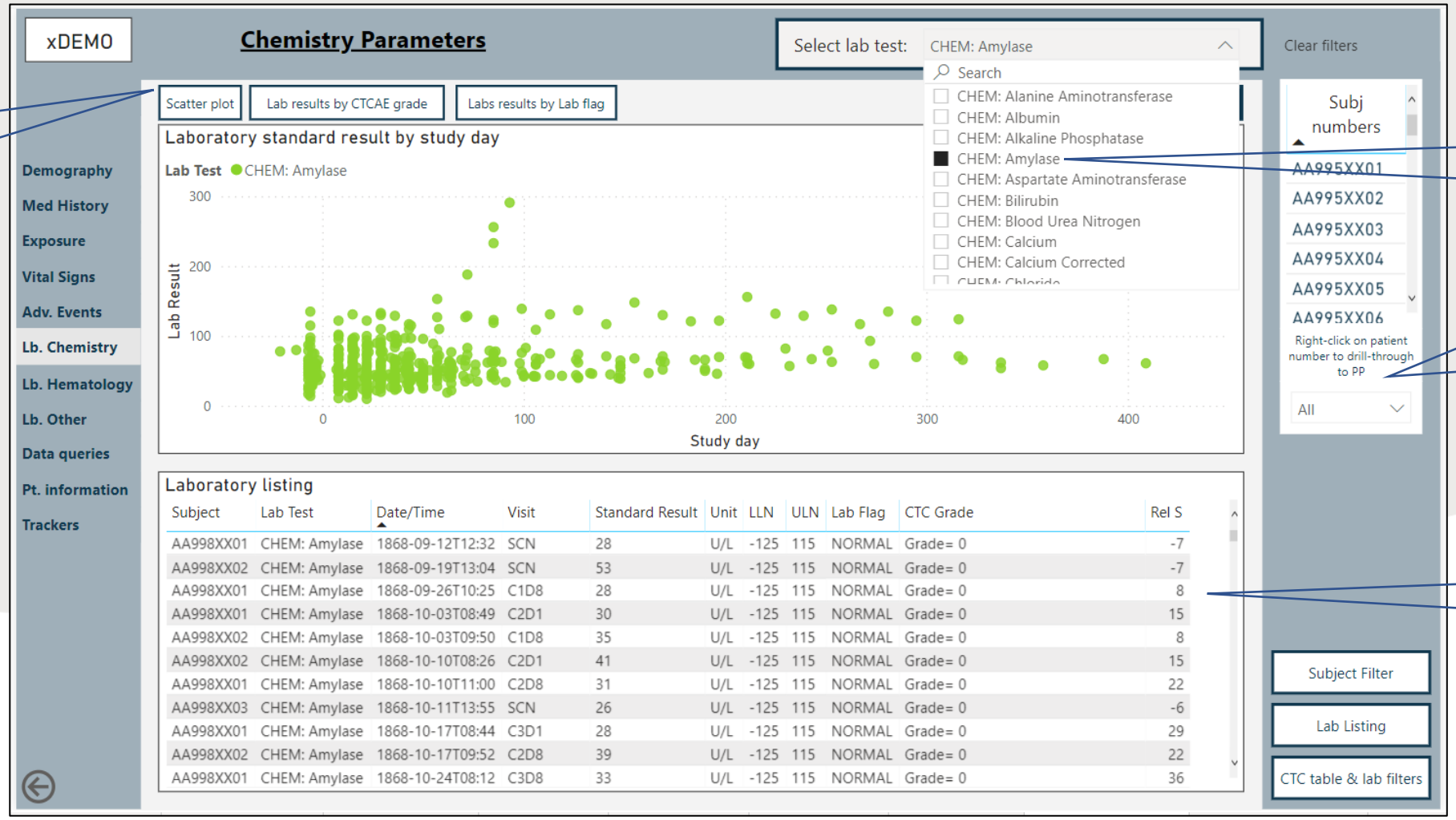
Switch to filter mode

Decomposition tree visual



Decomposition tree

Lab Data Visuals



Subject	Lab Test	Date/Time	Visit	Standard Result	Unit	LLN	ULN	Lab Flag	CTC Grade	Rel S
AA998XX01	CHEM: Amylase	1868-09-12T12:32	SCN	28	U/L	-125	115	NORMAL	Grade= 0	-7
AA998XX02	CHEM: Amylase	1868-09-19T13:04	SCN	53	U/L	-125	115	NORMAL	Grade= 0	-7
AA998XX01	CHEM: Amylase	1868-09-26T10:25	C1D8	28	U/L	-125	115	NORMAL	Grade= 0	8
AA998XX01	CHEM: Amylase	1868-10-03T08:49	C2D1	30	U/L	-125	115	NORMAL	Grade= 0	15
AA998XX02	CHEM: Amylase	1868-10-03T09:50	C1D8	35	U/L	-125	115	NORMAL	Grade= 0	8
AA998XX02	CHEM: Amylase	1868-10-10T08:26	C2D1	41	U/L	-125	115	NORMAL	Grade= 0	15
AA998XX01	CHEM: Amylase	1868-10-10T11:00	C2D8	31	U/L	-125	115	NORMAL	Grade= 0	22
AA998XX03	CHEM: Amylase	1868-10-11T13:55	SCN	26	U/L	-125	115	NORMAL	Grade= 0	-6
AA998XX01	CHEM: Amylase	1868-10-17T08:44	C3D1	28	U/L	-125	115	NORMAL	Grade= 0	29
AA998XX02	CHEM: Amylase	1868-10-17T09:52	C2D8	39	U/L	-125	115	NORMAL	Grade= 0	22
AA998XX01	CHEM: Amylase	1868-10-24T08:12	C3D8	33	U/L	-125	115	NORMAL	Grade= 0	36

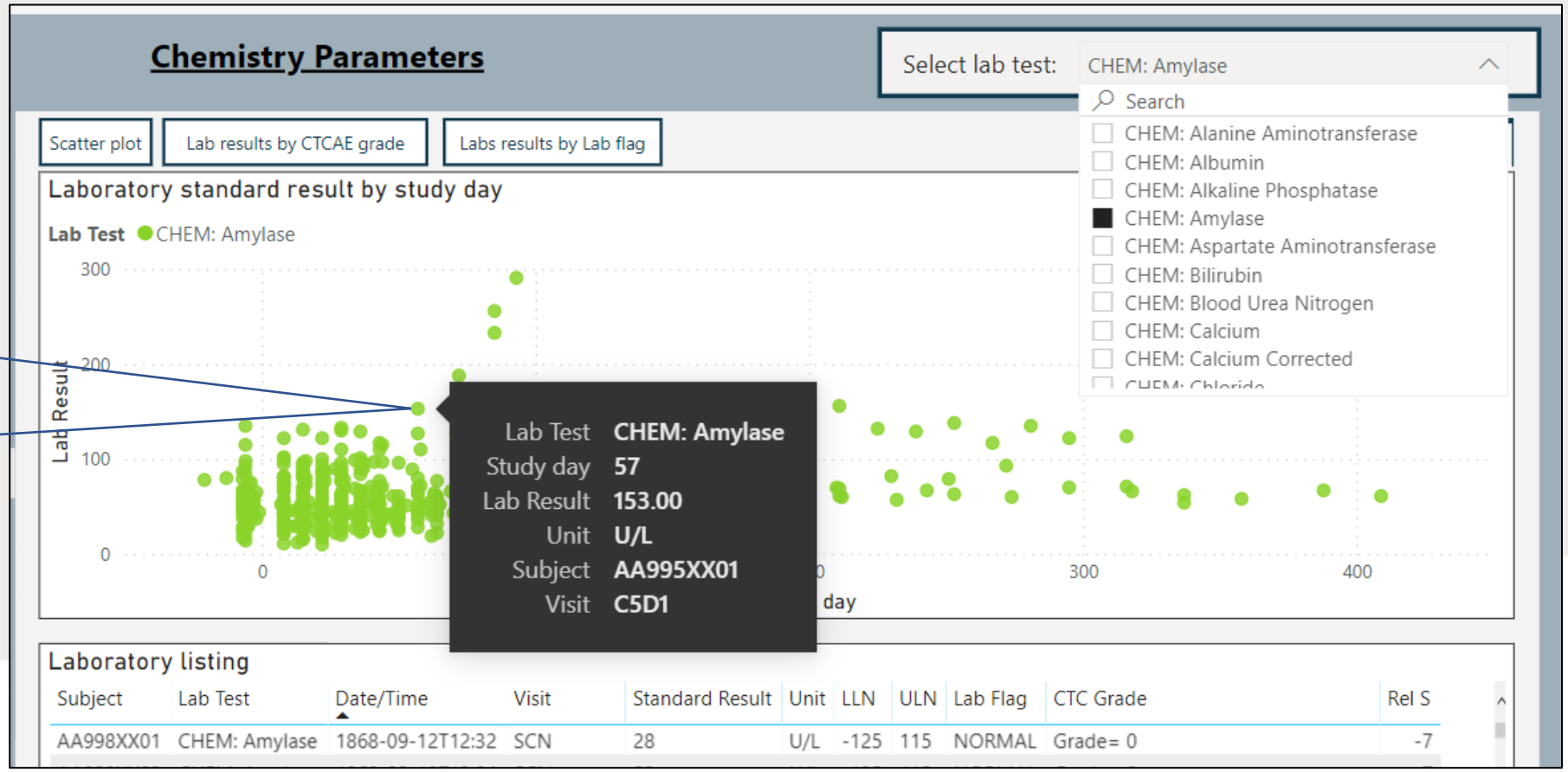
Change Chart Types

Select one, multiple or all tests

Drill-through to see complete patient profile

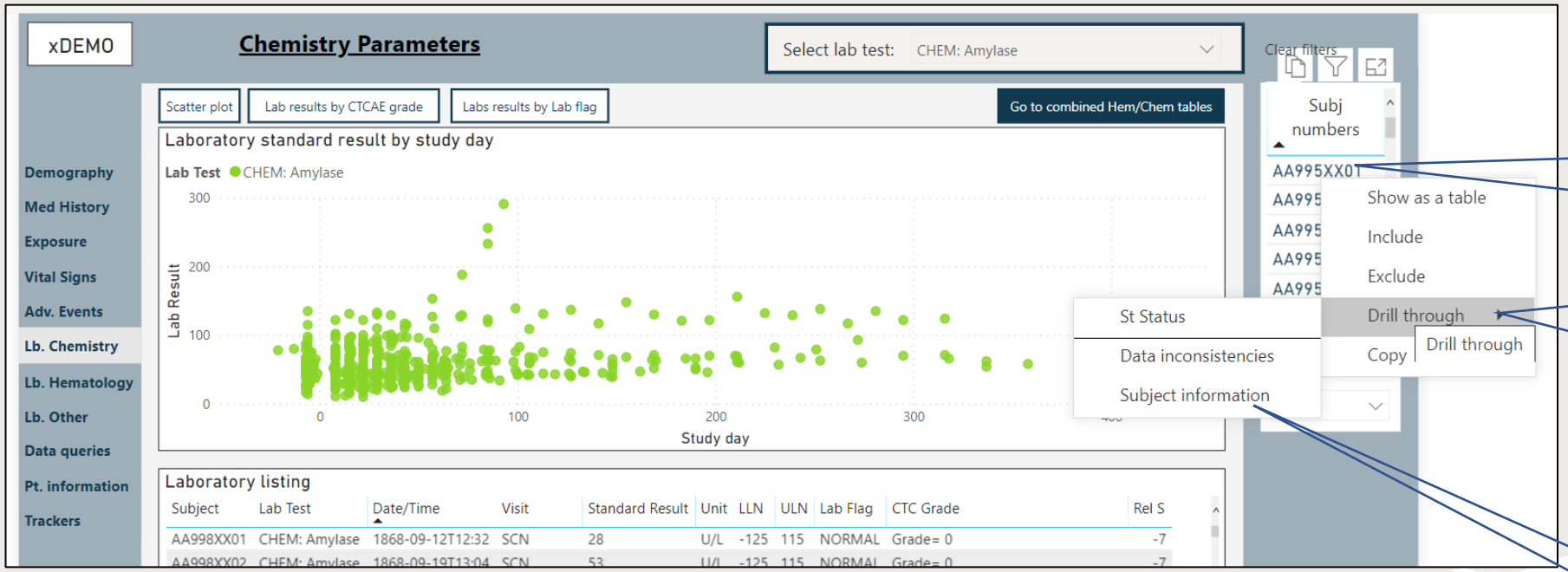
Subject listing

Identifying the subjects



Hover mouse pointer on individual points to see more details in the tooltip

Drill-through a subject

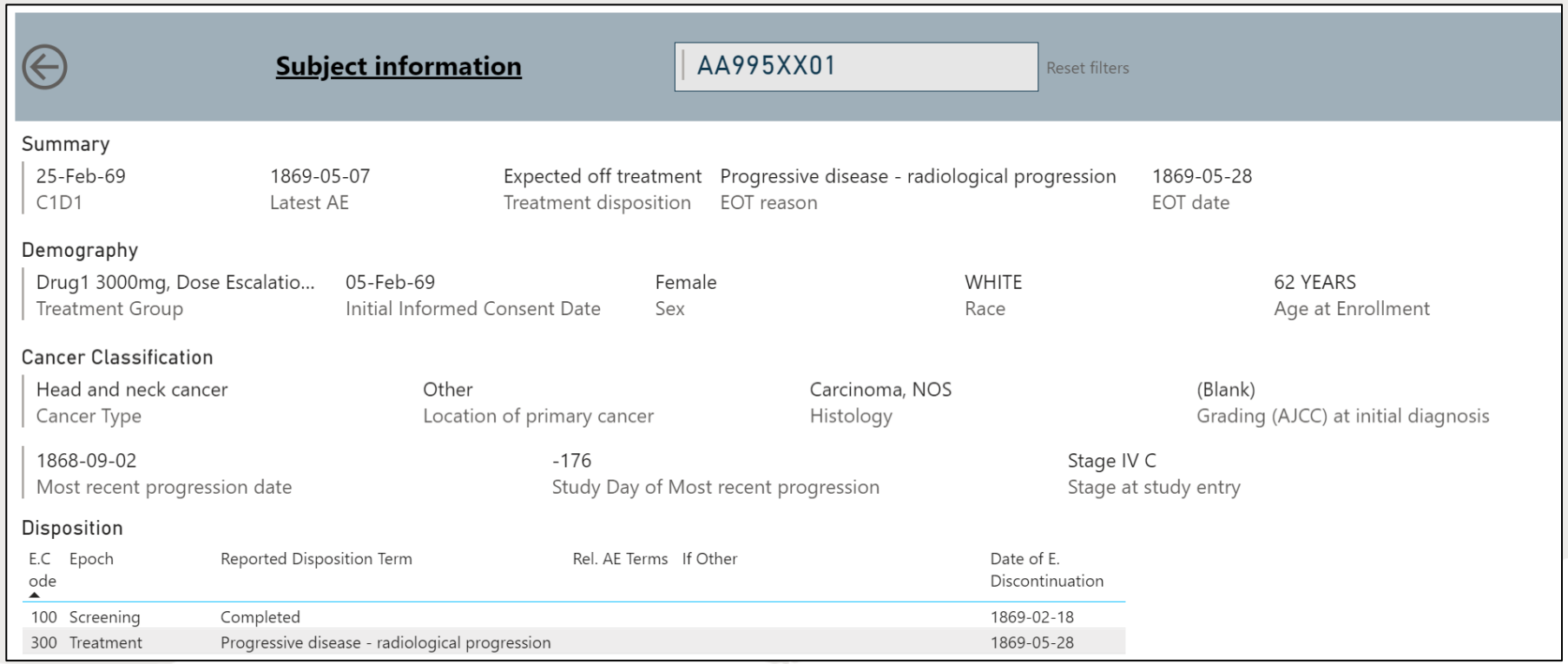


Right click on a subject

Select Drill through

Click on Subject Information

Subject Information (1/3)



Subject information AA995XX01 Reset filters

Summary

25-Feb-69	1869-05-07	Expected off treatment	Progressive disease - radiological progression	1869-05-28
C1D1	Latest AE	Treatment disposition	EOT reason	EOT date

Demography

Drug1 3000mg, Dose Escalatio...	05-Feb-69	Female	WHITE	62 YEARS
Treatment Group	Initial Informed Consent Date	Sex	Race	Age at Enrollment

Cancer Classification

Head and neck cancer	Other	Carcinoma, NOS	(Blank)
Cancer Type	Location of primary cancer	Histology	Grading (AJCC) at initial diagnosis

1868-09-02	-176	Stage IV C
Most recent progression date	Study Day of Most recent progression	Stage at study entry

Disposition

E.C ode	Epoch	Reported Disposition Term	Rel. AE Terms	If Other	Date of E. Discontinuation
100	Screening	Completed			1869-02-18
300	Treatment	Progressive disease - radiological progression			1869-05-28

Subject Information (2/3)

Death information

1870-01-01	Progressive disease	(Blank)
Date/Time	Primary Cause of Death	If Other, specify

Screening Characteristics

Test	Standard Result	Standard Unit
Performance Status	1 - Restricted active	
Heart rate	75	beats/min
Respiration rate	16	breaths/min
Temperature	36.9	C
Weight	82.6	kg
Diastolic blood pressure	88	mmHg
Systolic blood pressure	152	mmHg

Medical History

MH No	Reported Term	Start Date/Time	End Date/Time	Ongoing?	CTCAE grade at baseline	CTCAE not gradable
011	Heart Murmur	1806		Ongoing	1	
008	History of Shingles	1859	1859			Yes
006	Obesity	1865		Ongoing	1	
007	Low back pain	1865		Ongoing	2	
014	Facial Edema	1865-03		Ongoing	2	
013	Left facial pain	1865-03-11		Ongoing		Yes
003	Acquired hypothyroidism	1865-04		Ongoing	2	
004	Arthritis	1865-04		Ongoing	2	
005	Hyperlipidemia	1865-04		Ongoing	2	

Exposure

Drug1 3000mg, Dose Escalation (MTD)

Visit	SD Taken?	Actual Treatment	Dose	Regimen	Route	Start Date	Start Time	End Date	End Time	D. modification?	Reason	Rel. AE 1	Rel. AE 2	Rel. AE 3	Reason, Oth
Cycle 01 Day 01	Yes	Drugotinib 1	3000	Once every 2 weeks	Intravenous	25-Feb-69	11:30	25-Feb-69	12:44	None					
Cycle 02 Day 01	Yes	Drugotinib 1	3000	Once every 2 weeks	Intravenous	11-Mar-69	12:45	11-Mar-69	13:45	None					
Cycle 03 Day 01	Yes	Drugotinib 1	3000	Once every 2 weeks	Intravenous	25-Mar-69	12:47	25-Mar-69	13:55	None					
Cycle 04 Day 01	Yes	Drugotinib 1	3000	Once every 2 weeks	Intravenous	08-Apr-69	11:36	08-Apr-69	12:36	None					
Cycle 05 Day 01	Yes	Drugotinib 1	3000	Once every 2 weeks	Intravenous	22-Apr-69	11:21	22-Apr-69	12:28	None					

Subject Information (3/3)

Adverse events

Adverse Events

AE No	Reported Term	MedDRA PT	TEAE?	Start Date/Time	End Date/Time	Rel S/E	Rel to SD?	Rel
1	Right Shoulder Pain	Arthralgia	N	1869-02-18		-7/ -91	No	
2	decreased Lymphocyte count	Lymphocyte count decreased	Y	1869-03-11		15/ -70	No	
3	Intermittent left eye neuropathic pain	Neuralgia	Y	1869-03-19		23/ -62	No	
4	Discomfort to the left posterior roof of mouth	Oral discomfort	Y	1869-03-25		29/ -56	No	
5	Left cheek erythema close to left buccal wound	Erythema	Y	1869-03-26		30/ -55	No	
7	Sore throat	Oropharyngeal pain	Y	1869-03-29	1869-05-07	33/ -52	No	
10	Left cheek swelling	Swelling face	Y	1869-04-07		42/ -43	No	
11	Left cheek redness	Erythema	Y	1869-04-07		42/ -43	No	
8	Left facial pus drainage	Purulent discharge	Y	1869-04-22	1869-05-07	57/ -28	No	
9	Cancer related pain	Cancer pain	Y	1869-04-22		57/ -28	No	
12	Bilateral Leg Edema	Oedema peripheral	Y	1869-05-07		72/ -13	No	

Vital Signs

Vital Signs

Date/Time	Visit Name	Test	Standard Result	Standard Un
1869-02-18	SCN	Performance Status	1 - Restricted active	
1869-02-18T11:21	SCN	Diastolic blood pressure	88	mmHg
1869-02-18T11:21	SCN	Heart rate	75	beats/min
1869-02-18T11:21	SCN	Respiration rate	16	breaths/min
1869-02-18T11:21	SCN	Systolic blood pressure	152	mmHg
1869-02-18T11:21	SCN	Temperature	36.9	C
1869-02-18T11:21	SCN	Weight	82.6	kg
1869-02-25	C1D1	Performance Status	1 - Restricted active	
1869-02-25T11:23	C1D1	Diastolic blood pressure	78	mmHg
1869-02-25T11:23	C1D1	Heart rate	74	beats/min
1869-02-25T11:23	C1D1	Oxygen saturation	97	%
1869-02-25T11:23	C1D1	Respiration rate	16	breaths/min
1869-02-25T11:23	C1D1	Systolic blood pressure	148	mmHg
1869-02-25T11:23	C1D1	Temperature	36.9	C
1869-02-25T11:23	C1D1	Weight	81.6	kg

Systolic Blood Pressure: 93 Min, 160 Max, 129.20 Average

Diastolic Blood Pressure: 53 Min, 88 Max, 69.72 Average

Heart Rate: 73 Min, 85 Max, 77.64 Average

Respiration Rate: 16 Min, 20 Max, 17.54 Average

Temperature: 37 Min, 37 Max, 37.00 Average

Oxygen Saturation

Select vital signs parameters:
- Diastolic blood pressure
- Heart rate
- Oxygen saturation
- Respiration rate
- Systolic blood pressure
- Temperature
- Weight
- Performance Status

Vital Signs

ECG Findings and Cardiac Function

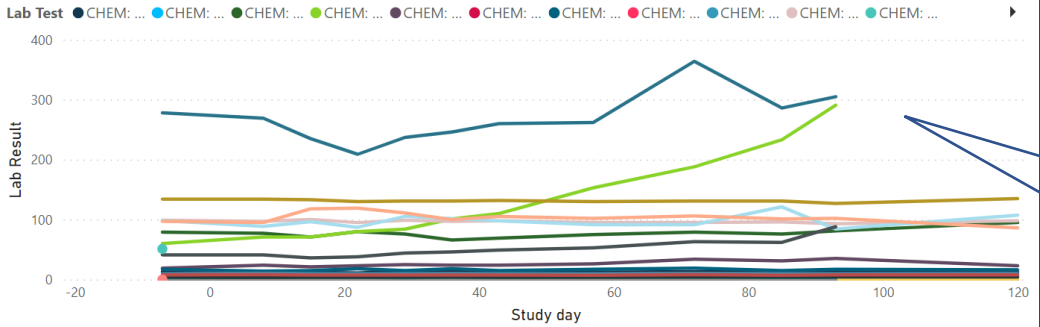
ECG Findings

Date/Time	Visit Name	Result/Finding	If Abnormal, Clinically Significant?	Link to Rave
1869-02-25T10:22	C1D1	NORMAL OR NORMAL VARIANT		www.google.com
1869-02-25T10:24	C1D1	NORMAL OR NORMAL VARIANT		www.google.com
1869-02-25T10:26	C1D1	NORMAL OR NORMAL VARIANT		www.google.com
1869-02-25T13:09	C1D1	NORMAL OR NORMAL VARIANT		www.google.com
1869-02-25T13:11	C1D1	NORMAL OR NORMAL VARIANT		www.google.com

Cardiac Function

Date	Visit name	Method	Results	Unit	Interpretation	If Abnormal, are findings clinically significant?
1869-02-18	SCN	Echocardiography	65.5	%	Normal or normal variant	
1869-04-01	C4D1	Echocardiography	70.4	%	Normal or normal variant	
1869-05-17	EOT	Echocardiography	65	%	Normal or normal variant	

Chemistry Graph: Value x Study Day



Laboratory

Lab Test	Date/Time	Visit name	Standard Result	Standard Unit	LLN	ULN	Lab Flag	CTC Grade
CHEM: Alanine Aminotransferase	1869-02-18T12:35	SCN	13	U/L	-150	52	NORMAL	Grade= 0
CHEM: Albumin	1869-02-18T12:35	SCN	3.6	g/dL	-118	4.6	NORMAL	Grade= 0
CHEM: Alkaline Phosphatase	1869-02-18T12:35	SCN	79	U/L	-150	140	NORMAL	Grade= 0
CHEM: Amylase	1869-02-18T12:35	SCN	60	U/L	-150	100	NORMAL	Grade= 0

Chemistry Graph

And a lot more

Conclusion

- Quickly visualize clinical trials data
- Draw quick insights
- Create Data Diagnostics System

- Create data transformation pipelines
- Create highly customized visualizations
- Process, summarize and visualize large datasets
- Create web applications with dashboards
- An alternative to R Shiny apps
- jupyter notebook - an alternative to R markdown file
- Create Clinical-Trial-At-a-Glance report after database lock in jupyter notebook

References

- Most popular programming language video - https://www.youtube.com/watch?v=ijjgPCDglLQ
- Covid-19 data visualizations - https://ourworldindata.org/grapher/covid-19-death-rate-vs-population-density?country=USA~GBR~CHN~RUS~BGD~CAN~AUS~HKG~HTI~OWID_WRL~IND~MCO~SGP~NZL~PAK~COG~NGA~MNG~VEN
- Power BI desktop download link - https://aka.ms/pbidesktopstore
- Microsoft documentation on Power BI - https://powerbi.microsoft.com/en-us/learning/
- Anaconda (python) download link - https://repo.anaconda.com/archive/Anaconda3-2021.05-Windows-x86_64.exe
- Plotly library for interactive graphs - https://plotly.com/python/
- PHUSE Open Data Repository - https://github.com/phuse-org/PODR/tree/master/sample_code

Thank you!

Questions/comments?