

PMSA 2023 Fall Symposium

Can ChatGPT truly do Serious Data Analytics?

JP Tsang, PhD (AI) & MBA (INSEAD)

President of Bayer

Xiaohao Ji

Director of Commercial Analytics, Kyowa Kirin



pmsa

PHARMACEUTICAL MANAGEMENT
SCIENCE ASSOCIATION

NOVEMBER 9-10, 2023

FLAMINGO HOTEL

LAS VEGAS, NV

Speakers



JP Tsang, PhD & MBA (INSEAD)
President of Bayser

Jean-Patrick Tsang is the Founder and President of Bayser, a Chicago-based consulting firm dedicated to sales and marketing for pharmaceutical companies. JP is an expert in data strategy and advanced analytics. JP has published 25+ papers, given 80+ talks at conferences, and completed 250+ projects. In a previous life, JP deployed Artificial Intelligence to automate the design of payloads for satellites. JP earned a Ph.D. in Artificial Intelligence from Grenoble University, advised 2 PhD students, and earned an MBA from INSEAD in Fontainebleau, France. He was the recipient of the 2015 PMSA Lifetime Achievement Award.



Xiaohao Ji,
Director Commercial
Analytics, Kyowa Kirin

Xiaohao Ji is a seasoned leader in commercial analytics and insights, with over 17 years of experience in the pharmaceutical and consulting sectors. Xiaohao is specialized in providing fact-based, actionable brand insights and diagnosing business health through story-telling KPIs and deep-dives. He has an unwavering passion to harnessing ever-evolving technology to generate faster, better and cheaper business insights. Xiaohao holds a Master of Science in Engineering from University of Pennsylvania and a Bachelor of Science from Tsinghua University in China.

Abstract

ChatGPT has taken the world by storm and everyday someone is claiming yet another feat that ChatGPT has pulled off. ChatGPT now comes in as an API and has been endowed with countless plug-ins including Wolfram Alpha for heavy lifting in math. At the same time, we are also witnessing a mushrooming of LLM models.

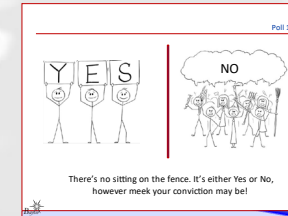
The question is obvious and compelling: “Can ChatGPT truly do serious data analytics?” The answer to this question has huge and transformative implications for our industry.

ChatGPT is very powerful but not without shortcomings. It hallucinates and makes up very convincing but wrong answers. It has seen a drop in IQ since inception. Also, it is sometimes fickle. There are days it responds to a prompt and another when it refuses to answer the same prompt.

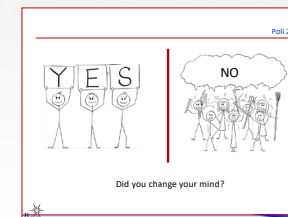
The way we explored this question is through experimentation. We looked at several Use Cases in Commercial Analytics, provided the requisite data, and asked the analysis question. In this talk, we’ll describe the experiments, present our findings, and reflect on their implications.

Agenda

- I. The Question
- II. First Reaction
- III. On Second Thoughts
- IV. Word on Data X Ray
- V. Our Crystal Ball
- VI. Q & A



Poll 1



Poll 2

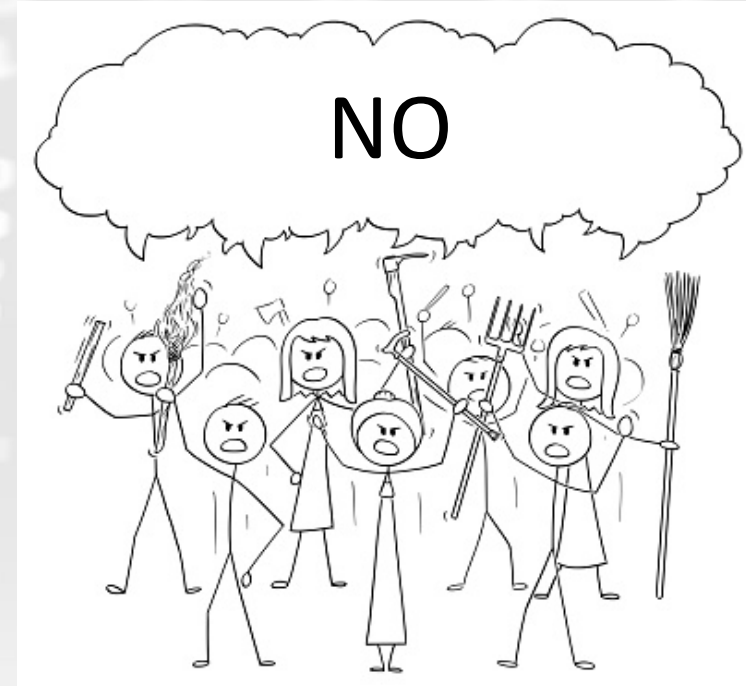
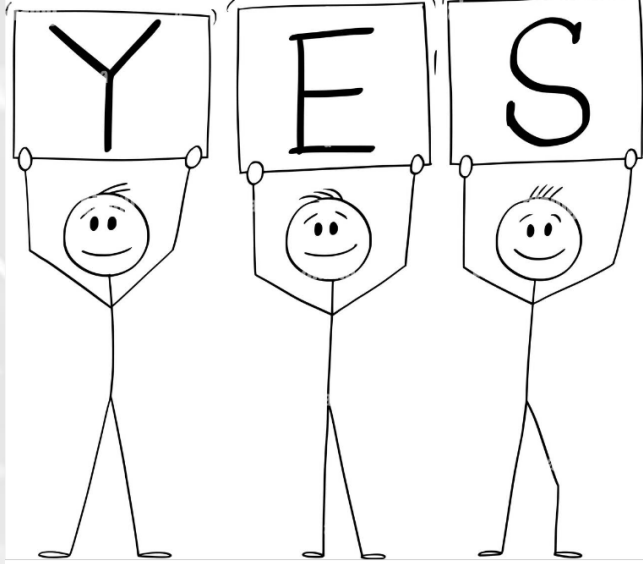
I. The Question

Key Question



**Can ChatGPT
Truly do Serious
Data Analytics?**

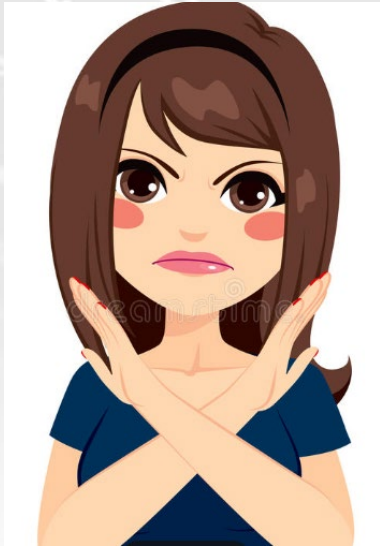
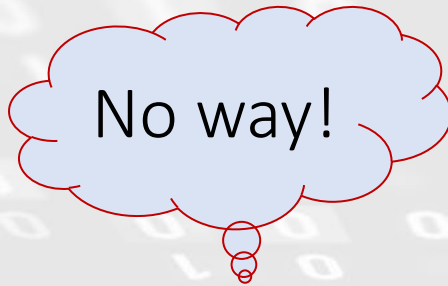
Poll 1



There's no sitting on the fence. It's either Yes or No, however meek your conviction may be!

II. First Reaction

First Reaction – No Way!



Can ChatGPT truly do
Serious Data Analytics?

ChatGPT or any LLM for that matter ...

1. **Hallucinates** – Makes things up out of thin air. Blunders go undetected as they sound compelling but are false.
2. **Fickle** – One day one answer, the following day another answer!
3. **Sick Leave** – Under the weather some days and needs to take a day off!
4. **Not up to date** – The most recent data used for training is as of Jan 2022 for ChatGPT 3.5 and Apr 2023 for ChatGPT 4.0.
5. **Does not know how to count!**

ChatGPT cannot count words in a list!

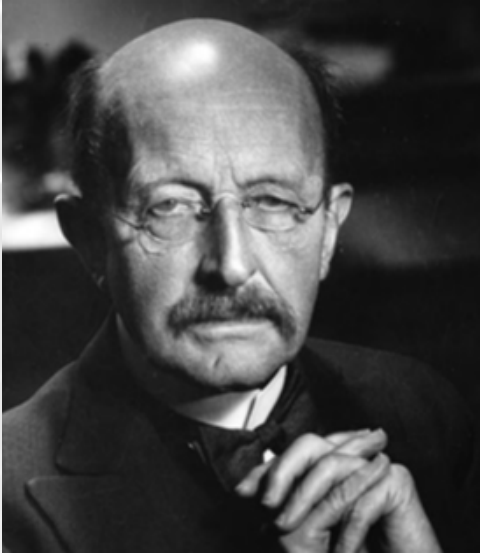
Nb	Answer – ChatGPT4.0
1	There are 658 words in the list provided.
2	The total number of words in the list you provided is 858.
3	The list contains 900 words.
4	There are 1,003 words in the list provided.
5	The list contains 1000 words.
6	The total number of words in the given list is 801.
7	There are 923 words in the list given.
8	There are 857 words in the list given.
9	The total number of words in the list is 844.
10	There are 650 words in the list provided.

Above the responses to the same counting prompt.

The correct answer is 1000.

III. On Second Thoughts

Food for Thought



“When you change the way
you look at things, the things
you look at change.”

— **Max Planck**

Max Karl Ernst Ludwig Planck was a German theoretical physicist whose discovery of energy quanta won him the Nobel Prize in Physics in 1918. Planck made many substantial contributions to theoretical physics, but his fame as a physicist rests primarily on his role as the originator of quantum theory.

Our Thesis



While the objections against ChatGPT are valid, there are workarounds.

If you heed our advice, you'll be amazed by the analytics ChatGPT can do for you.

7 Ingredients in the Secret Sauce



The Secret Sauce to Leveraging
ChatGPT for Serious Data Analytics

1. Prompt Engineering
2. Break Down in Baby Steps
3. Use Plug-ins to Access the Larger World
4. Judicious Use – Invoke ChatGPT when there is no simple and direct way to get the answer.
5. Panel of LLMs – Use not one but several LLMs to get the best answer (remember the ML Ensemble Technique?)
6. Workhorse data source – Make sure the data source is the right one for the job
7. Leverage Multiple External Sources to provide perspective on what the data captures

Secret Sauce 1 – Prompt Engineering

1. Always give ChatGPT an identity.
2. Provide a clear objective and relevant context. Be as specific as possible.
3. Describe precisely the format of the response you are expecting (e.g., various fields of a table).
4. Define constraints for the response (e.g., number of words, in what style, max number of words, not to make things up, etc.)

Prompt Engineering – Generic Prompt

Scenario: Understanding the competitive landscape of a specific therapeutic area

ChatGPT Prompt without Prompt Engineering:

“Provide information on competitive landscape for treatment of NSCLC with EGFR mutation”

First-Generation EGFR TKIs:

1. **Gefitinib (Iressa):** One of the first EGFR TKIs (tyrosine kinase inhibitors) approved for treating EGFR-positive NSCLC.
2. **Erlotinib (Tarceva):** Another first-generation inhibitor widely used in treating NSCLC with EGFR mutations.

Second-Generation EGFR TKIs:

1. **Afatinib (Gilotrif):** Known for its irreversible binding, Afatinib has shown better efficacy in some cases compared to first-generation TKIs.
2. **Dacomitinib (Vizimpro):** Offers longer progression-free survival than earlier TKIs in some studies but comes with more adverse effects.

Third-Generation EGFR TKIs:

1. **Osimertinib (Tagrisso):** The current leader in the third-generation category, Tagrisso is approved for first-line treatment and is effective against T790M mutations, which are responsible for resistance against first and second-generation TKIs.

Upcoming Treatments:

- Several other targeted therapies and combination therapies are in clinical trials.

Checkpoint Inhibitors:

- Immune checkpoint inhibitors like Pembrolizumab (Keytruda) and Nivolumab (Opdivo) are also being explored in combination with EGFR TKIs.

Prompt Engineering – Optimized Prompt



ChatGPT Prompt with Structured Prompt Engineering:

- Act as an expert in market research and competitive intelligence in the US market.**
- Your task is to help understand the competitive landscape of the FDA approved treatment for NSCLC with EGFR mutation.**
- Your communication style is professional and data-driven.**
- Include the information such as: Product name, Manufacturer, Molecule, Drug Class, Approval Date, Mode of Action**
- Present the information in tabular format with each row representing a product**
- Do not make up any information!”**

Product Name	Manufacturer	Molecule	Drug Class	Approved Indication	Approval Date
Afatinib (Gilotrif)	Boehringer Ingelheim Pharmaceuticals Inc.	Afatinib	Tyrosine Kinase Inhibitor	Metastatic NSCLC with EGFR exon 19 deletions or exon 21 (L858R) substitution mutations	2013 ^{1 2}
Dacomitinib (VIZIMPRO)	Pfizer Pharmaceutical Company	Dacomitinib	Tyrosine Kinase Inhibitor	First-line treatment of metastatic NSCLC	Sept. 27, 2018 ^{3 4}
Erlotinib (Tarceva)	Genentech, Inc.	Erlotinib	Tyrosine Kinase Inhibitor	Treatment of NSCLC	November 18, 2004 ^{5 6 7}
Gefitinib (Iressa)	AstraZeneca UK Limited	Gefitinib	Tyrosine Kinase Inhibitor	Metastatic NSCLC with EGFR exon 19 deletions or exon 21 (L858R) substitution mutations	July 13, 2015 ⁸
Mobocertinib (Exkivity)	Takeda Pharmaceuticals, Inc.	Mobocertinib	Tyrosine Kinase Inhibitor	Treatment of locally advanced or metastatic NSCLC	September 15, 2021 ⁹
Osimertinib (TAGRISSO)	AstraZeneca Pharmaceuticals LP	Osimertinib	Tyrosine Kinase Inhibitor	Adjuvant therapy after tumor resection in NSCLC patients	December 18, 2020 ¹⁰
Rybrevant (amivantamab-vmjw)	Janssen Pharmaceutical Companies of Johnson & Johnson	Amivantamab-vmjw	Bispecific antibody	Treatment of NSCLC with EGFR exon 20 insertion mutations	Not specified ^{11 12}
Amivantamab-vmjw	Not specified	Amivantamab-vmjw	Bispecific antibody	Treatment of locally advanced or metastatic NSCLC with EGFR Ex20ins mutations	May 21, 2021 ¹³

Prompt Engineering – Minimize hallucination

Use "Don't make things up" constraint to minimize hallucination



 what's the annual sales in the US for Xtandi, Zytiga, Erleada and Nubeqa in 2020? Present the information in tabular format
< 3 / 3 > 


 As of my last update in January 2022, the most recent complete annual sales data I have is from 2020. Please note that the numbers are approximations and may vary depending on the source.

Here is the annual sales data for Xtandi, Zytiga, Erleada, and Nubeqa in the United States for the year 2020:

Drug	2020 Annual Sales in the U.S. (USD)
Xtandi	~ \$1.6 billion
Zytiga	~ \$1.1 billion
Erleada	~ \$350 million
Nubeqa	~ \$60 million

Keep in mind that these numbers can be influenced by various factors such as changes in medical guidelines, availability of generic versions, and competition from other medications.

 What's the annual sales in the US for Xtandi, Zytiga, Erleada and Nubeqa in 2020? Present the information in tabular format.
Don't make up any numbers if you don't know this information.
< 4 / 4 > 

 I don't have access to real-time or updated databases, and my training data only goes up until January 2022. However, I can say that annual sales for pharmaceutical products like Xtandi, Zytiga, Erleada, and Nubeqa would typically be reported in the financial statements of the companies that produce them or in market research reports. These reports are generally available to the public, although some may require a subscription or purchase. You can also find this kind of information through governmental databases that track pharmaceutical sales.

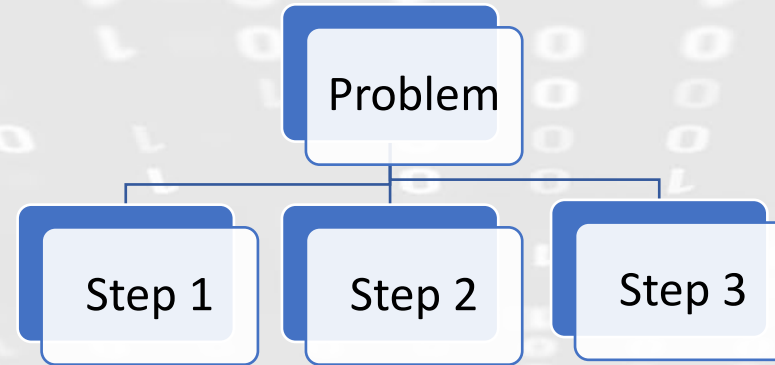
Note: The sales numbers provided on the left were not accurate. This demonstrates the importance of this prompt when asking for hard-to-find information

Secret Sauce 1 – Prompt Engineering

Additional Tips:

- For list-type queries, the response may not be complete or exhaustive. You may run the prompt a few times to cross-check.
- Don't ask for overly long and complex response, as it may cause error or incomplete results

Secret Sauce 2 – Break Down in Baby Steps



1. Takes away the meta-problem of breaking a problem into sub-problems, thereby reducing the odds of introducing mistakes stemming from problem decomposition
2. Significantly increase the computational attention dedicated to each sub-problem
3. Improve transparency of the process and minimize errors due to incorrect interpretation of the problem

Break Down in Baby Steps – Example – 1 of 4

Scenario: Conduct segmentation and clustering analysis for a group of HCPs, and evaluate the call effectiveness for each cluster.

Baby Step 1: Load and validate input data

- **Use this Noteable space:** <https://app.noteable.io/p/782425d1-4923-4d38/PMSA-Test>
- The file "Sample HCP Universe.xlsx" contains HCP-level database that should be used for this analysis. **Load and validate** the data and **provide key stats**.

Baby Step 2: Describe the objective, context and requirements

- **Act as an experienced market analyst.** Your goal is to create 3-5 segments from the HCP database to prioritize those with the highest potential for writing more product TRx in the near future.
- **Make succinct response** with key rationales and results. **Ask clarifying questions** if needed. Don't make things up.
- **Make sure to import appropriate libraries** in the python code.

Break Down in Baby Steps – Example – 2 of 4

Baby Step 3: Continue with promotional response analysis

- *The **dependent variable** is incremental product TRx ($\text{prod_trx_c12m} - \text{prod_trx_p12m}$).*
- *Create **a linear regression model** to analyze the coefficients of different components that drives incremental sales. These includes a intercept component, a component related to historical TRx volume (prod_trx_p12m), and a component related to $\ln(1 + \text{rep_calls_c12m})$.*
- *The **output should be a table** with the value of these coefficients.*
- *Create **a response curve chart** where x axis is the number of calls, and y axis is the incremental TRx driven by the calls.*

Baby Step 4: Stratify the response curves by segment

- *Create **a second model** that stratify the coefficients by Tier value (High/Medium/Low). List out the coefficients in a table along with the $P > |t|$ value. **Plot the three response curves** on the same chart.*

Break Down in Baby Steps – ChatGPT Responses

Missing Values

- No missing values were found in the dataset.

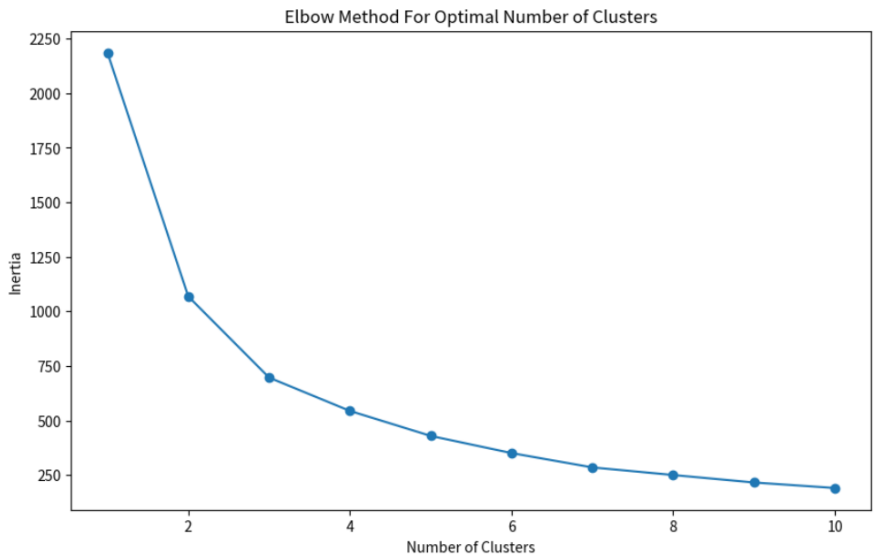
Summary Statistics

- The dataset contains 1,093 HCPs.
- Key metrics like ``Prod_TRx_C12M``, ``Mkt_TRx_C12M``, ``Prod_TRx_P12M``, and ``Mkt_TRx_P12M`` have varying ranges and standard deviations, which will be important for segmentation.

Data Types

- The dataset contains a mix of numerical and categorical variables.

With the data validated, I'll proceed with the segmentation analysis. The primary metrics for segmentation will be historical market and product volume (``Prod_TRx_C12M`` and ``Mkt_TRx_C12M``) in the recent 12 months. Secondary metrics like ``Specialty`` and ``Market_Decile`` will also be considered if necessary.

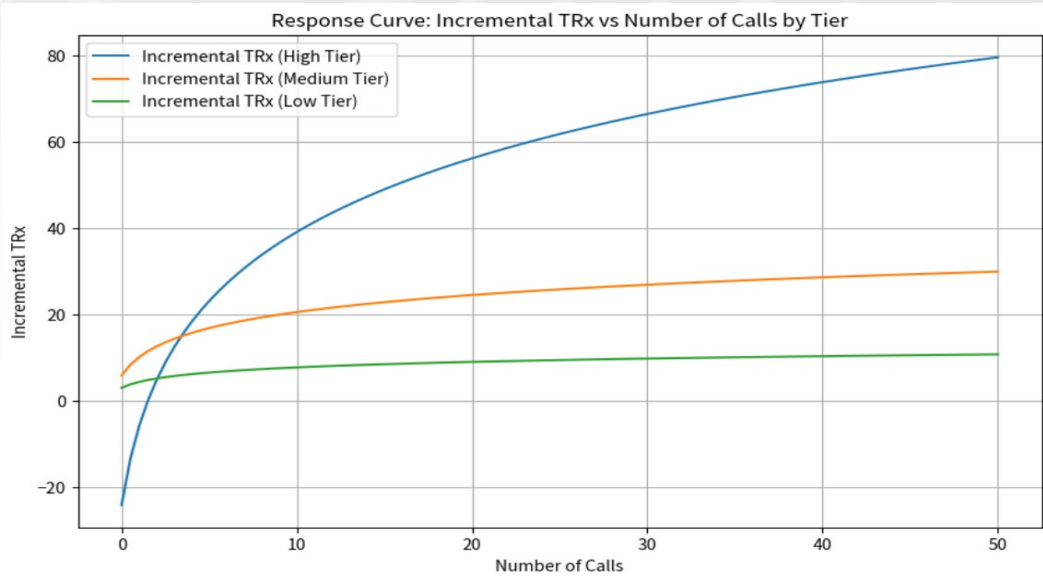
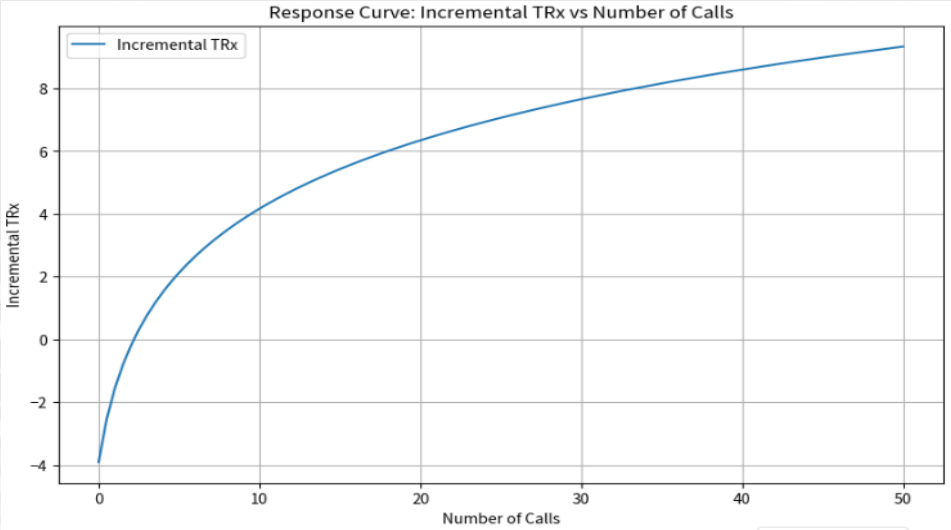
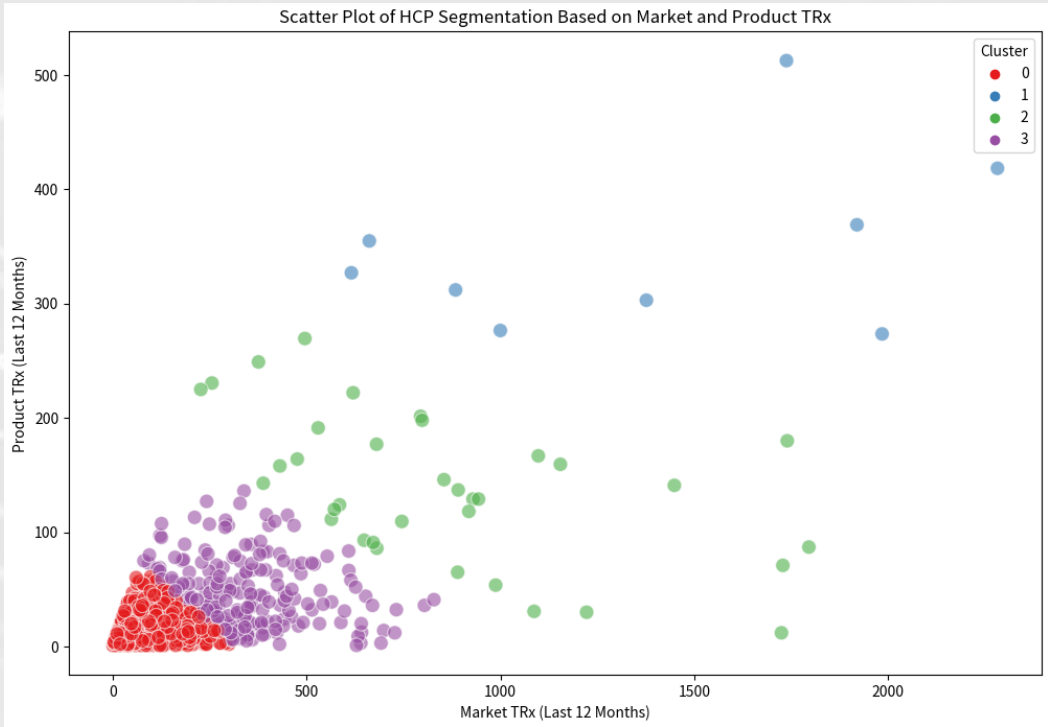


Cluster	Number_of_HCPs	Avg_Prod_TRx_C12M	Std_Prod_TRx_C12M	Avg_Mkt_TRx_C12M	Std_Mkt_TRx_C12M
1	9	349.659	76.6101	1384.66	621.423
2	35	137.695	63.5027	855.76	424.382
3	190	48.5451	29.5175	353.298	146.782
0	859	14.2114	11.41	73.4025	61.6138

- **Cluster 1:** Highest priority with 9 HCPs, an average Product TRx of 349.659, and an average Market TRx of 1384.66.
- **Cluster 2:** Second priority with 35 HCPs, an average Product TRx of 137.695, and an average Market TRx of 855.76.
- **Cluster 3:** Third priority with 190 HCPs, an average Product TRx of 48.5451, and an average Market TRx of 353.298.
- **Cluster 0:** Lowest priority with 859 HCPs, an average Product TRx of 14.2114, and an average Market TRx of 73.4025.

The table also includes the standard deviation for Product and Market TRx within each cluster, which can be useful for understanding the variability within each segment.

Break Down in Baby Steps – ChatGPT Responses



Coefficients Table

- Intercept (const): -3.9055
- Historical TRx Volume (Prod_TRx_P12M): -0.1354
- Natural Logarithm of Rep Calls (ln_Rep_Calls_C12M): 3.3667

Break Down in Baby Steps – Additional Notes

- To use the "Noteable" plugin, structured data needs to be first uploaded to the Noteable.io website.
- Always verify if ChatGPT correctly translated the request into Python code.
- Do not submit sensitive data if you don't trust the security/privacy policies of the plugins

Secret Sauce 3 – Plug-Ins

- 900+ Plug-ins significantly expand ChatGPT's capability beyond the confines of the LLM model.
- Better execute specific use cases that ChatGPT is not optimized for.
- Connection to other cloud services can significantly improve productivity in some cases

Note:

Exercise caution when using plugins. There's added risk for security and privacy as it's difficult to verify author credentials

Plug-Ins – Examples



Noteable

Uninstall (x)

Create notebooks in Python, SQL, and Markdown to explore data, visualize, and share notebooks with everyone.

Developer info  



Wolfram

Uninstall (x)

Access computation, math, curated knowledge & real-time data through Wolfram|Alpha and Wolfram Language.

Developer info  

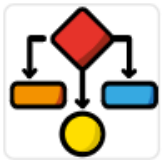


Zapier

Uninstall (x)

Interact with over 5,000+ apps like Google Sheets, Gmail, HubSpot, Salesforce, and thousands more.

Developer info  



Diagrams: Show Me

Uninstall (x)

Schemes, Diagrams, Architecture Visualisations, Flow-Charts, Mind Maps. Export and Edit for free!


Developer info  



Ai PDF

Uninstall (x)

Super-fast, interactive chats with PDFs of any size, complete with page references for fact checking.


Developer info  



Link Reader

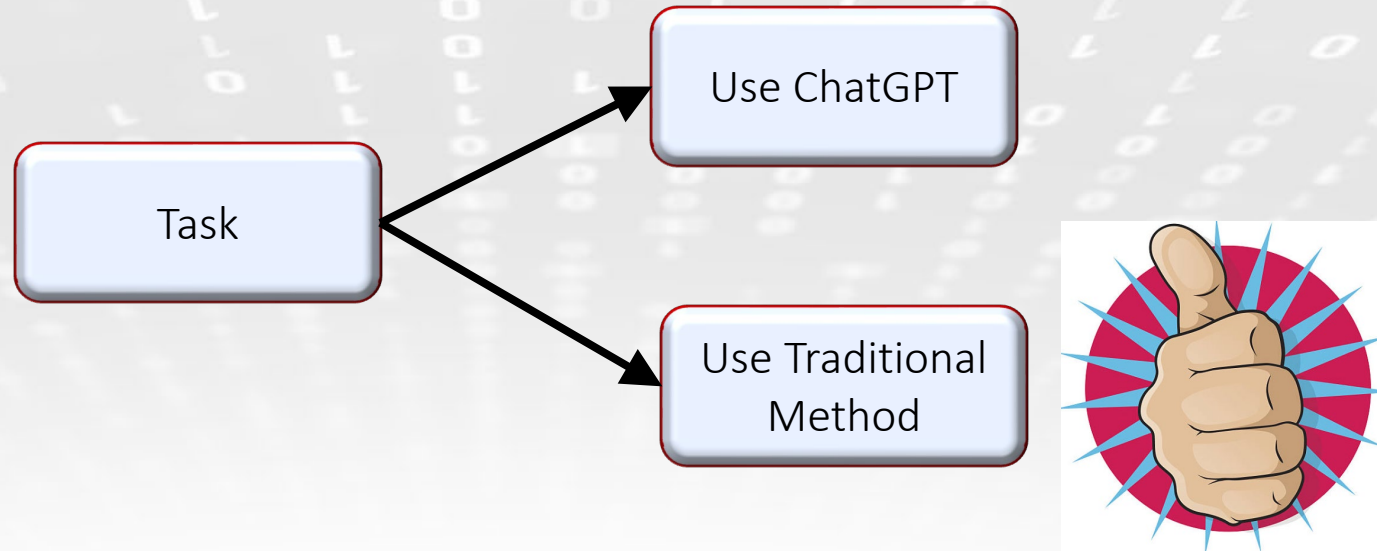
Uninstall (x)

Read any links: webpage, youtube, PDF, PPT, image, Word doc etc; OCR image; Search & browse web.

Developer info  

Secret Sauce 4 – Judicious Use

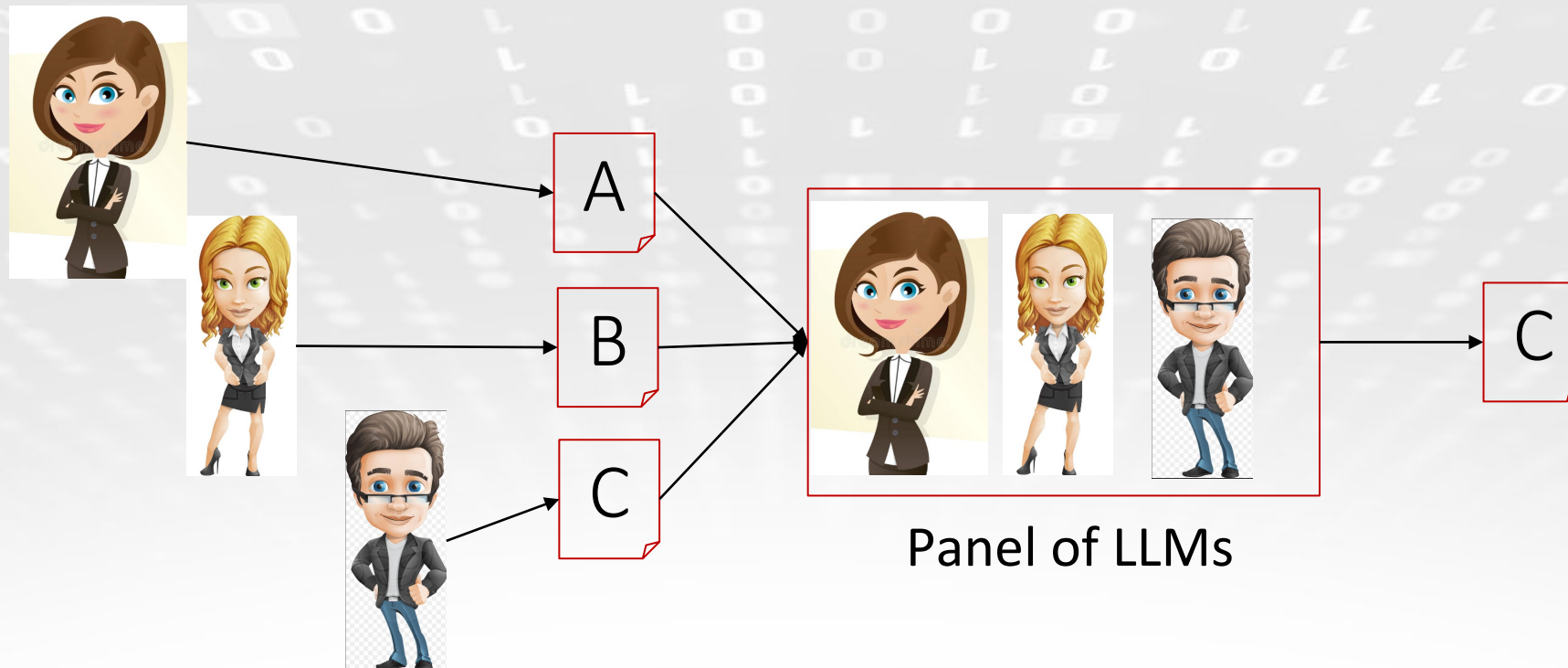
When there is a choice between using ChatGPT or the traditional method to perform a task, prefer the traditional method.



1. No hallucination
2. Same reliable answer all the time

Secret Sauce 5 – Panel of LLMs

Observation: For one task, the LLM you use may provide the best answer. For another task, the best answer may be from another LLM.



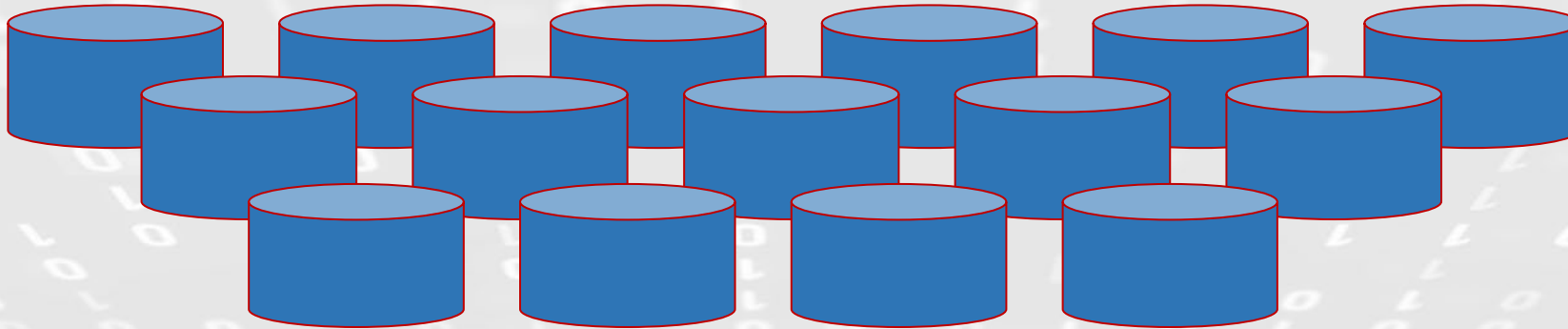
Approach: Ask a Panel of LLMs to choose the best answer from the answers provided by the LLMs on the Panel.

Secret Sauce 6 – Workhorse Data Source



1. Make sure the workhorse dataset that will be providing the analytical insights is the right one for the job, especially in regards to the Use Case and Drug/Therapeutic Area.
2. Use Data X Ray to gauge fitness of the workhorse dataset prior to running the analytics.

Secret Sauce 7 – Leverage External Data Sources



1. *Sick people*
2. *Afford therapy*

Demand



Supply

1. *Physicians*
2. *Hospitals/Beds*

Reference Data
Sources

1. *ICD*
2. *NDC*
3. *CPT*

IV. Word on Data X Ray

Data X Ray



*Assessing a data source is not only patient counts.
It is much more!*

Completeness

Granularity

Linkability

Accuracy

Timeliness

USP

Stability

Scorecard – Dimensions and Examples

Completeness	Geographic holes, poor capture of IDNs (e.g., FCS, Texas Oncology, RCCA, etc.), Payer under representation, missing interactions in patient journey etc.
Accuracy	Market Share upside down (data blocking, leakage) Patient, age, gender, and ethnicity bias.
Stability	Data Vendor lost a supplier
Granularity	Only Medicare, Medicaid, Commercial, and Other No BCBSL of IL Group-Bin-PCN.
Timeliness	Lag too big for alerts. Data span too short for patient journey
Linkability	Contains Datavant tokens to link to other data sources Ids are universal (e.g., NPI), internal (need lookup table), or encrypted (e.g. patient id)
USP	Contains NCPDP id of the Pharmacy. Paid amounts in addition to charged. Adjudication status

Key Questions

The database only gives the numerator

What is the denominator?

How biased or
representative?

Fit for the job?

What's the Use case? What's the Drug?

Data X Ray – Key Features

*External Data
Sources*

*Showstoppers
&
Setbacks*

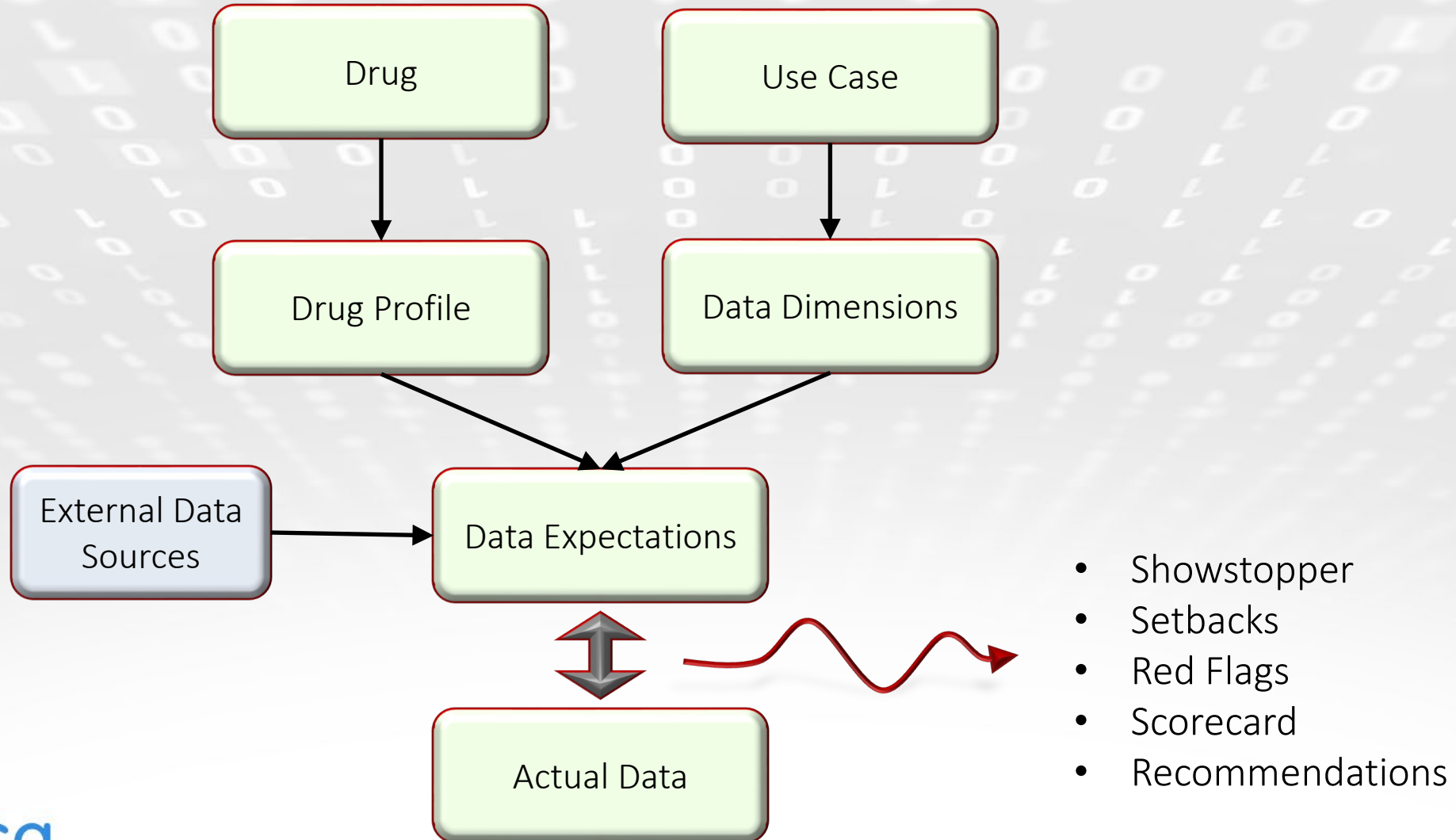
Scorecard

*Use Case
&
Drug/TA/DA*

*Annotates
Dictionary*

*Leverages
ChatGPT*

Key Idea



Examples of Showstoppers and Setbacks

Showstopper!

Physician Targeting

NPI is encrypted

Showstopper!

Buy and Bill IV

No J-code in the Mx data

Showstopper!

Multiple indications

Only 1 ICD-10 code in
Mx data

Setback!

Prostate Cancer

Poor representation of
Medicare

Setback!

Migraine is 3 times more
prevalent in women than
men

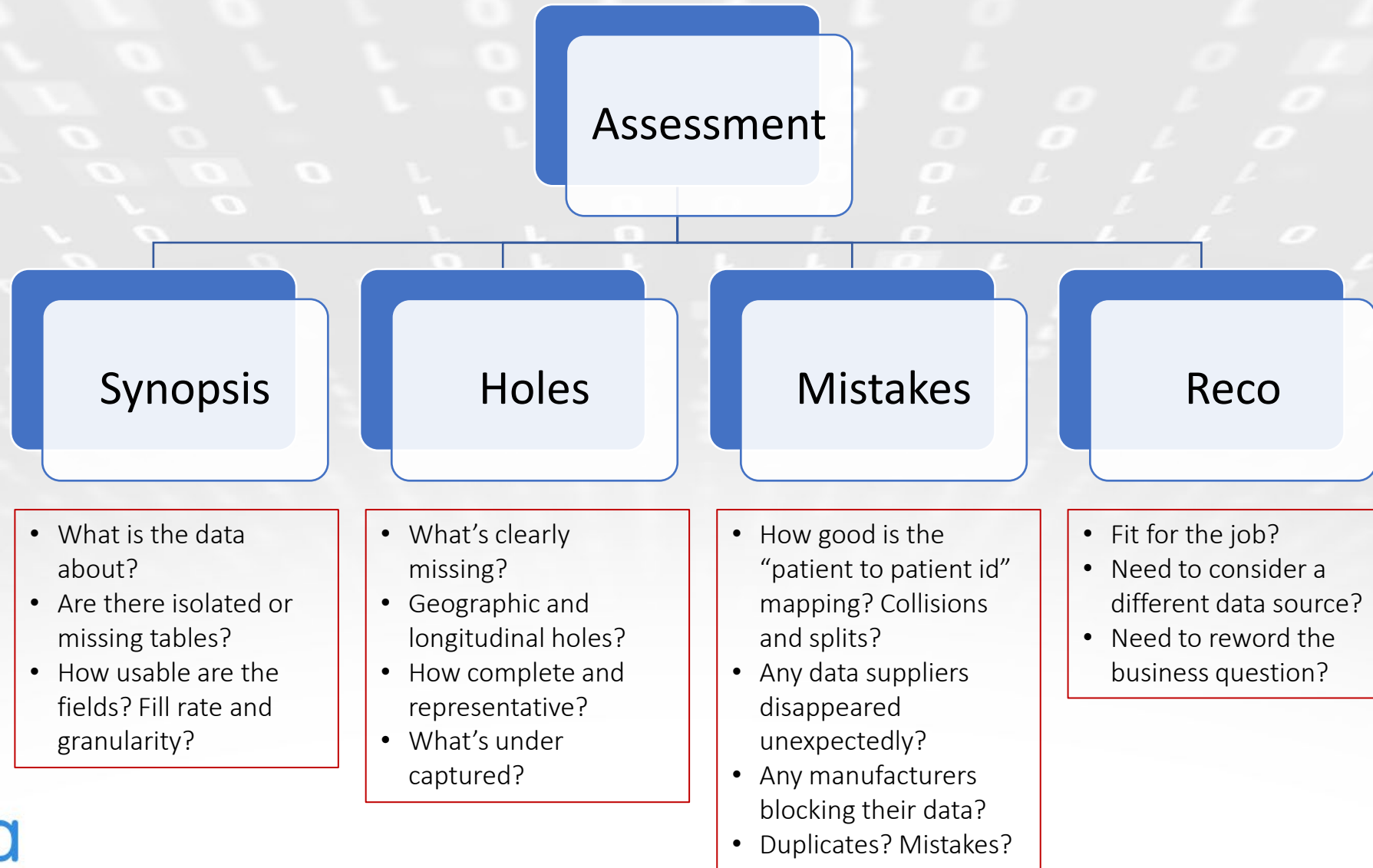
Men-Women: 50-50

Setback!

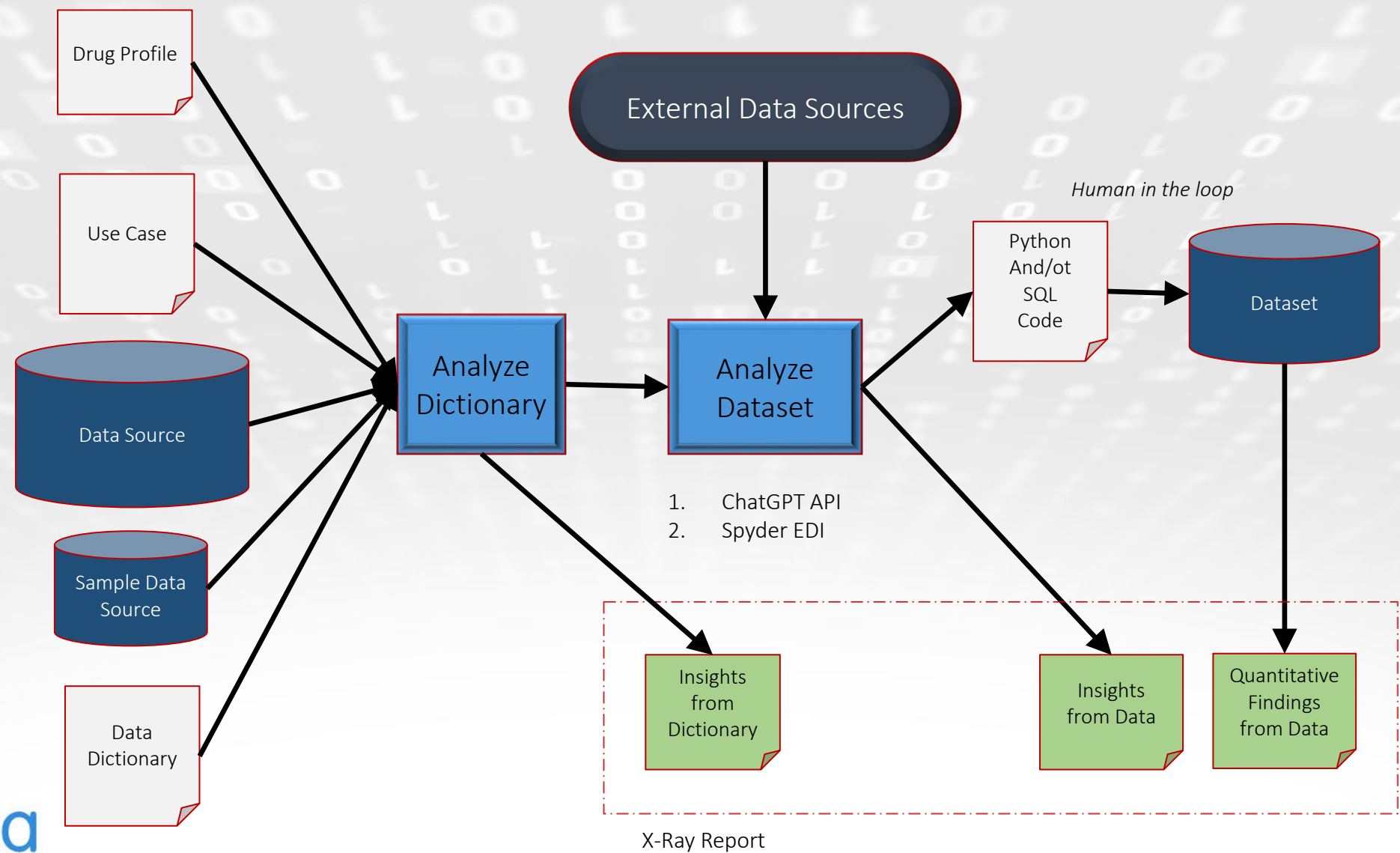
Companion Diagnostics

No CPT-4 code for lab
procedures
No lab results either

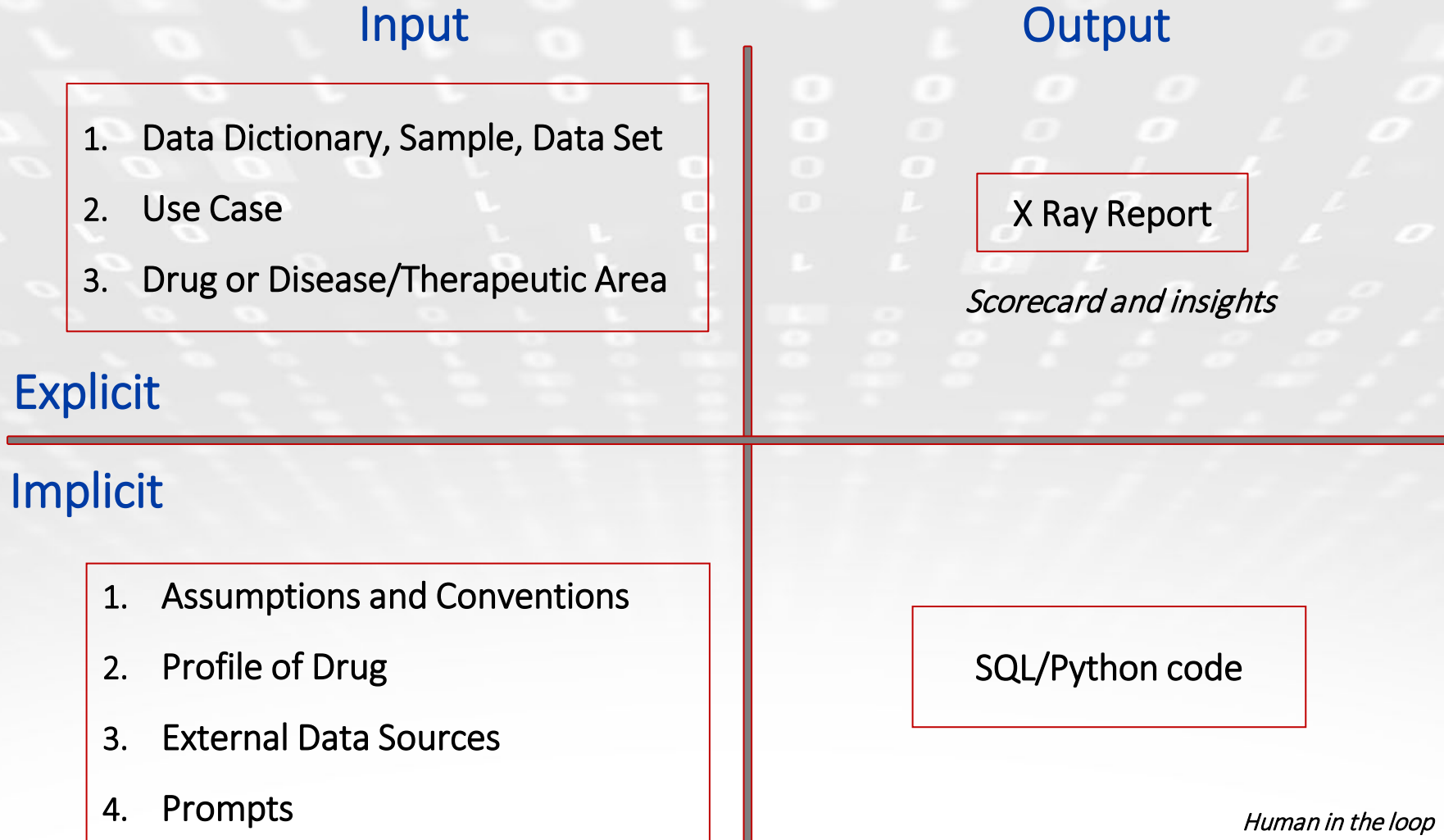
Data Source Assessment



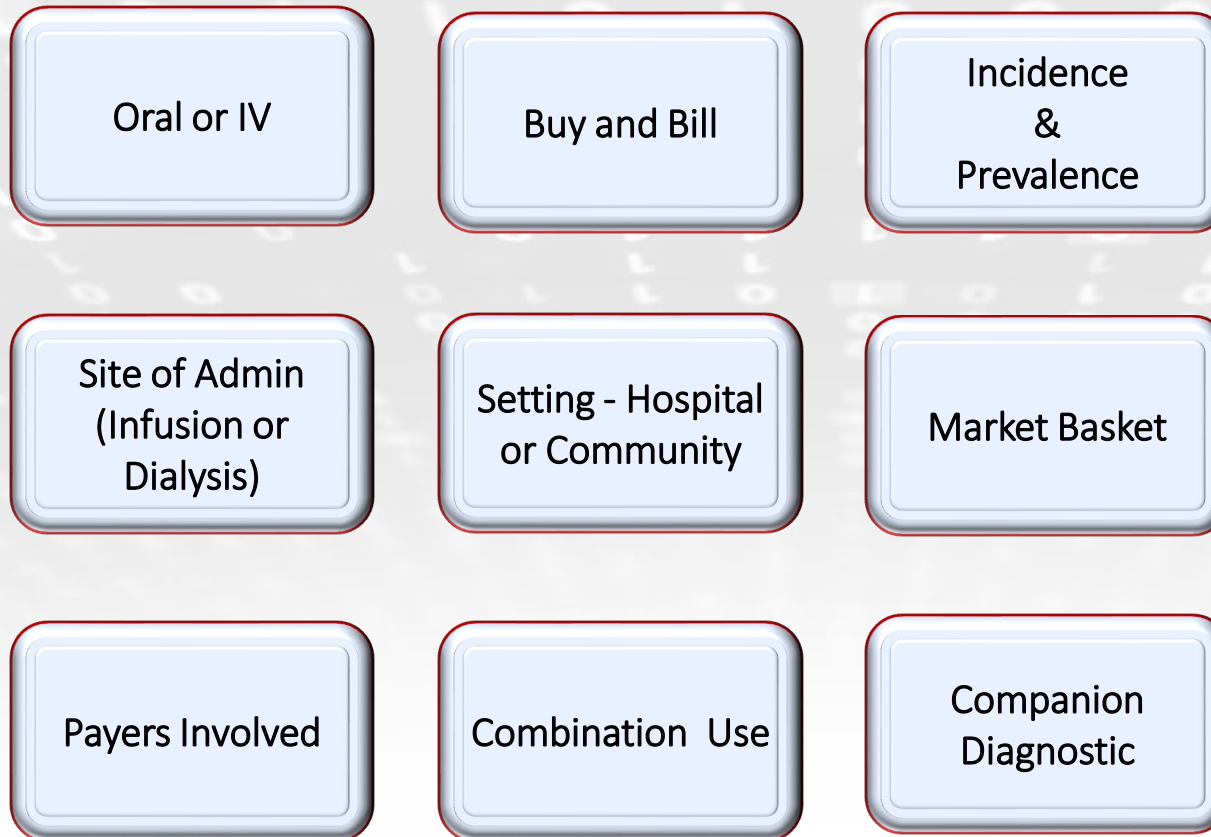
Architecture of the X-Ray Machine – Input and Output



Input-Output



Drug Profile – Key Dimensions



- TA/DA
- Indications
- Acute vs. Chronic vs. PRN
- Distribution (SP, SD, brick and mortar, etc.)
- Orphan Status
- REMS Status
- Lifecycle
- Survival Rate
- Etc.

REMS = Risk evaluation and mitigation strategies

Common Use Cases – Top 15



V. Our Crystal Ball

What does the Future Has in Store For Us

How is data analytics going to change in the coming years?



You ain't see nothing yet!
The upcoming changes will be more dramatic than what you've ever seen. More disruptive than DDD, Xponent, PLD, and EMR combined.

Key Point 1 – User Aware of Caveats and Shortcomings

- ❑ The user is made fully aware of the caveats and shortcomings of the workhorse data source prior to analysis.
- ❑ The user may decide to drop the dataset and use another one instead that is better suited for the use case and TA in question.
- ❑ Insights gleaned may be adjusted in light of the shortcomings and blind spots of the data source.

Key Point 2 – Democratization of Analytics

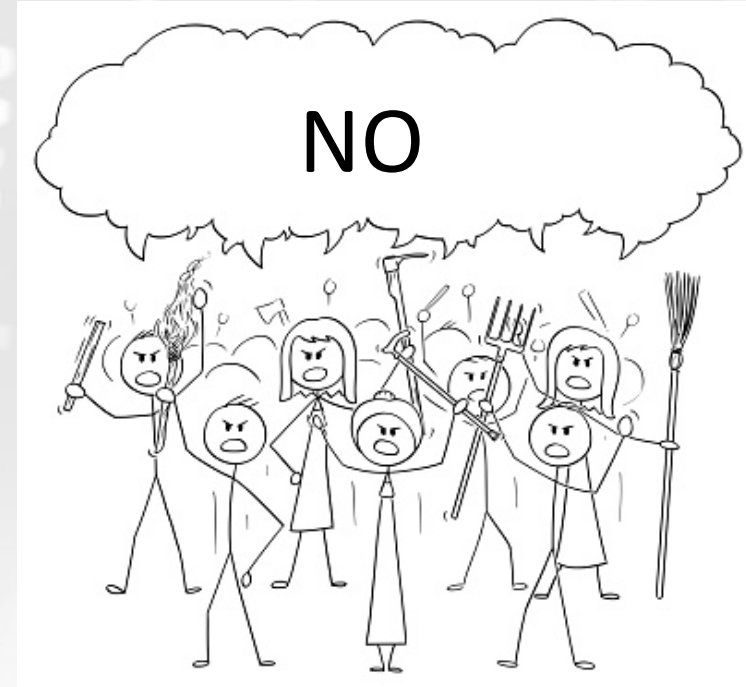
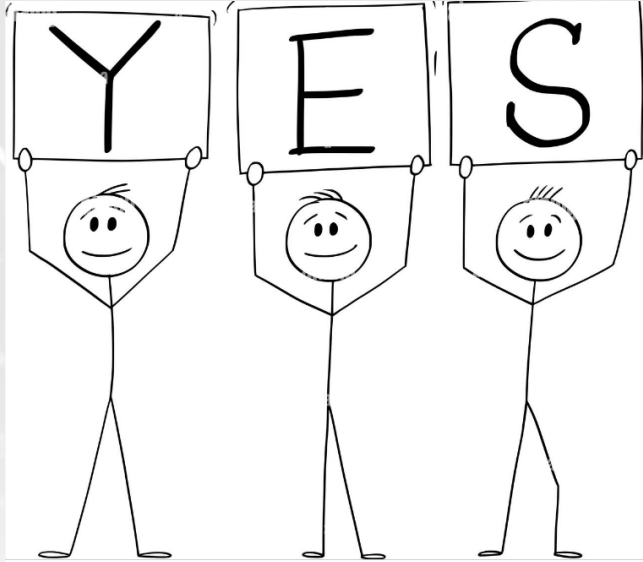
- ✓ Programming skills are no longer needed.
- ✓ Explosion in number of questions explored, profiles of answer seekers, and types of analyses performed.
- ✓ Broadening of insights gleaned.
- ✓ Deepening of our understanding of the potential and limitations of a data source to answer questions.

Key Point 3 – All kinds of Data Formats

- ❖ Data no longer needs to be in tabular form
- ❖ Can be unstructured text, images, videos, libraries of DNA, proteins, and the like.
- ❖ Enables a mind-boggling range of analyses never thought possible before!

Poll 2

Can ChatGPT truly do Serious Data Analytics?



Did you change your mind?

VI. Q & A

Q&A



Your Question

Please send questions to

JP – bayser@bayser.com or (847) 224 8086

Ji – jixiaohao@gmail.com (609) 902 1282