ADACTA

Podatkovna pismenost i moderni kontroler

Milan Listeš, Regionalni voditelj BI odjela | Hrvatska i Srbija

Why are we here today?

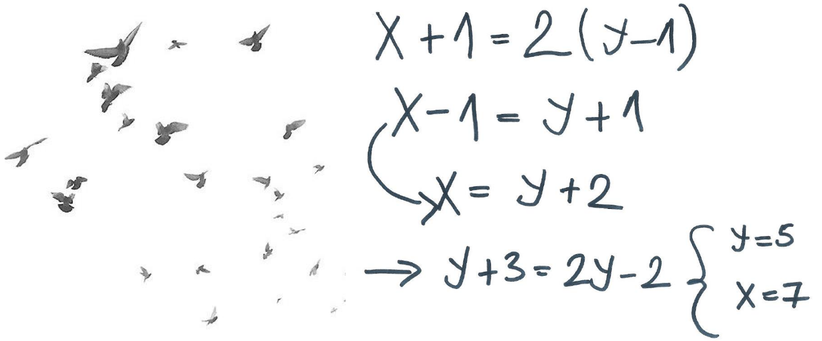
- Because we have a problem ☺
  - Modern faster technology
  - But same concepts
  - Just let me extract everything
  - But when I get a new question
- Many reasons for that > **data literacy**
- Why aren't we data literate?

What did we do for past 25y?

- We were good with data… But just few.

  - Peter has 7 apples. If he gives away 1 apple, how many apples will Peter have?

  - Two bird flocks are flying side by side. First tells the second: „Give us one bird and we'll have twice as more than you have", and second tells the first one: „You give us one bird and we'll have the same". How many birds are in each flock?



What is data literacy?

Science and technology [edit]

- 4G[58]
- Aggregator[63]
- Agile[64]
- Ajax[8][58]
- Algorithm[65]
- Benchmarking[66]
- Back-end[31]
- Beta[8]
- Big data - larger data sets than last month [67]
- Bleeding edge[31]
- Blog[63] – plus various other words that incorporate "blog"
- Bring your own Device - use of personal equipment (usually mobile) in a work environment
- Bricks-and-clicks[31][31]
- Clickthrough[31]
- Cloud[68]
- Collaboration[69]
- Content management[69]
- Content Management System[63] – also known as CMS.
- Convergence[70]
- Cross-platform[31]
- Cyber-physical Systems (CPS)
- Datafication[71]
- Data mining[72] - any kind of data collection or analysis, even simple statistics such as taking averages on large data sets
- Data science[73]
- Deep dive[17]
- Deep web[74] - used interchageably with "Dark web" even though they're not the same
- Design pattern[75]
- DevOps[76]
- Digital divide[63]
- Digital Remastering[77]
- Digital Rights Management[8] – also known as DRM.
- Digital signage[78]
- Disruptive Technologies[79]
- Document management[69]
- Dot-bomb[15][31]
- ____-Driven Development[80]
- E-learning[81]
- End-to-End
- Engine[82]
- Enterprise Content Management[63] – also known as ECM.
- Enterprise Service Bus[83] – also known as ESB.
- Evolution - Often use ambiguously in political or sociological arguments in reference to theories of social Darwinism. (e.x. "Society has evolved.")
- Framework[8]
- Folksonomy[63]
- Fuzzy logic[84]
- Growth Hacking
- HTML5[85]
- Immersion[86]
- Information superhighway / Information highway[8]
- Internet of Things[87]
- Innovation[88]
- Machine Learning
- Mashup[8]
- Microservices
- Mobile[89]
- Modularity[90][91]
- Nanotechnology[92]
- Netiquette[63]
- Next Generation[86] (also "NextGen")
- PaaS
- Podcasting[58][63]

Education [edit]

- Accountable talk[23]
- Higher-order thinking[24]
- Invested in[25]
- Run like a business[26]
- Student engagement[27][28][29]
- Common Core
- Bloom's Taxonomy
- Differentiated instruction
- Digital Literacy
- Flipped Classroom

„Data literacy includes the ability to **read, work with, analyze** and **argue with** data."*

**MIT & Emerson University*

Data literacy consists of…

Data fundamentals

- Understanding data
- Understanding aggregation
- Understanding distribution

Fundamental analytics

- Signal & noise
- Correlation and causation
- Confidence intervals
- Analytical a/b testing
- Linear regression
- Hypothesis testing
- Design of experiment

Data-informed decision making

- Introduction
- Framework
- Analytics techniques

Advanced analytics

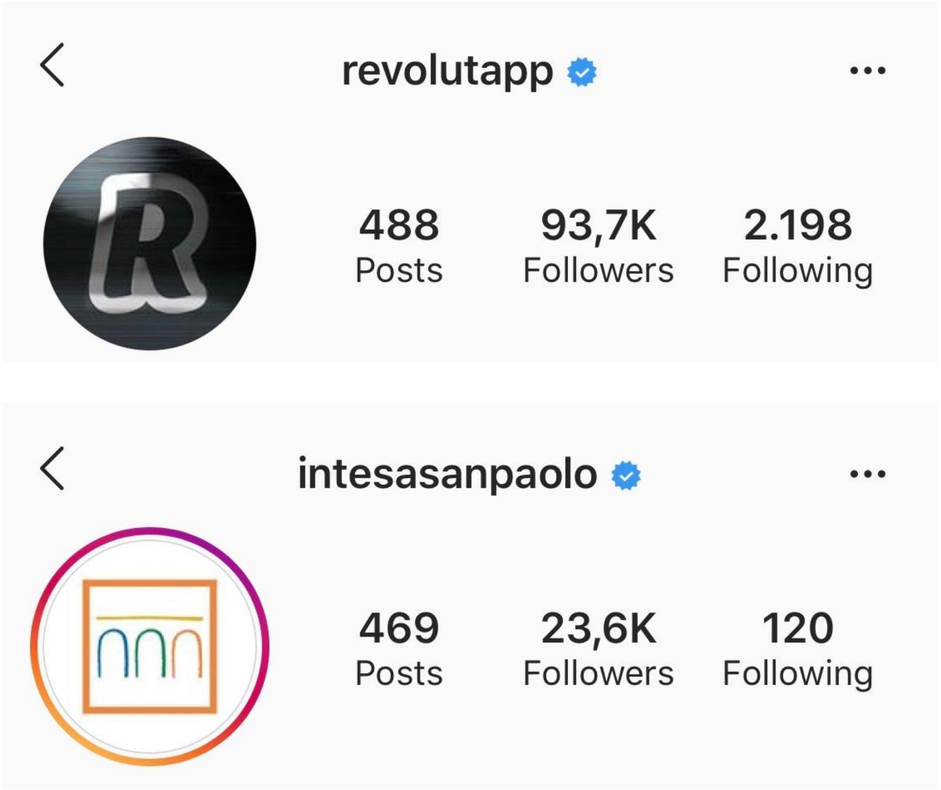
- Decision tree
- Advanced integration with QS
- Multiple Linear Regression Analysis

Data fundamentals

- Understanding data
- Understanding aggregation
- Understanding distribution

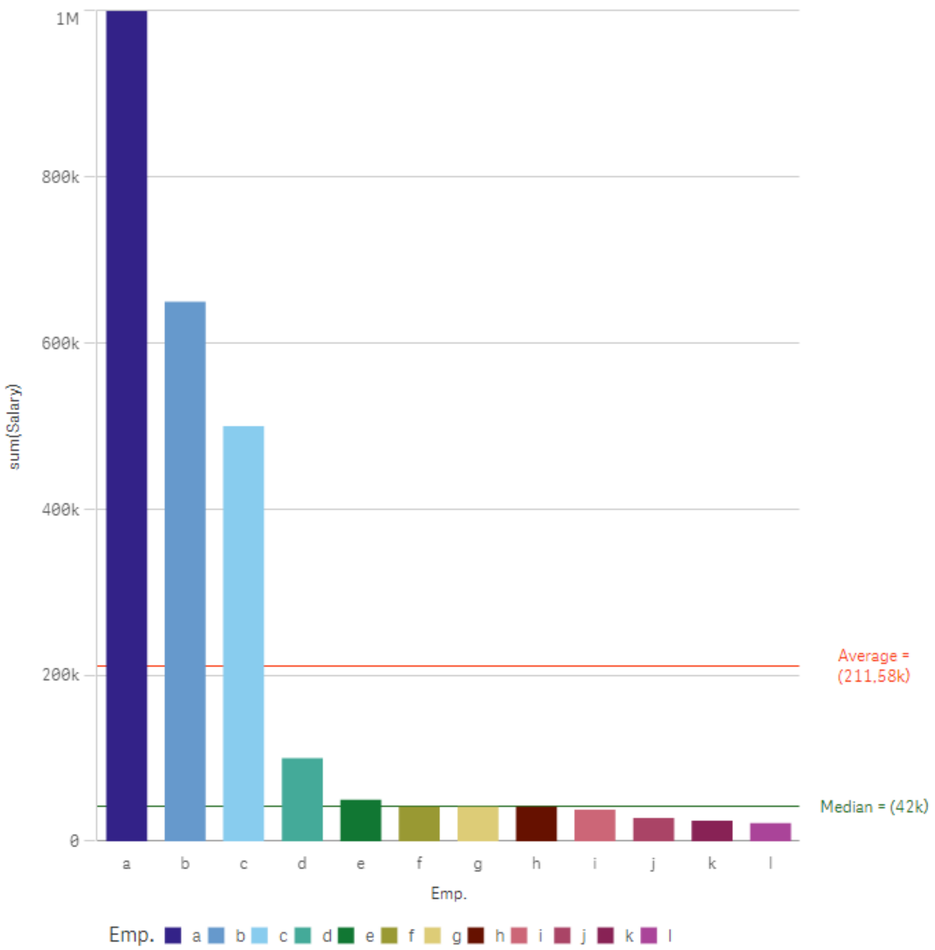
Data is not just…

- Data > Currency > Oil > bla bla bla > What does it mean?
- **Use case |** Hospitality / banking / insurance answering q who is my customer? > **customer segmentation**
  - Transactional data
    - **Hospitality |** reservations, spending, additional services
    - **Insurance |** premiums, claims, master data filled in the form
    - **Banking |** transactions, credit amounts/repayments, savings
  - Everything else
    - Social media for **product response**
    - Unstructured media for **sentiment analysis / opinions**
    - Wearables for **personalized profiling**
    - Etc…



Sum, Avg, what else?

- One of most misused functions business have seen ☺
  - Sum > OK
  - Avg > Not OK ☺
- **Use case |** Answering a question what is the average payroll in my company? > **new employee negotiation, annual budgeting, employer branding**
  - If we don't have a normal distribution, and we usually don't have in labour intense industries > **manufacturing, field service**
    - **Average <> Median**
    - **Difference** 169.000$
- Many more use cases that require this kind of interpretation



Data fundamentals

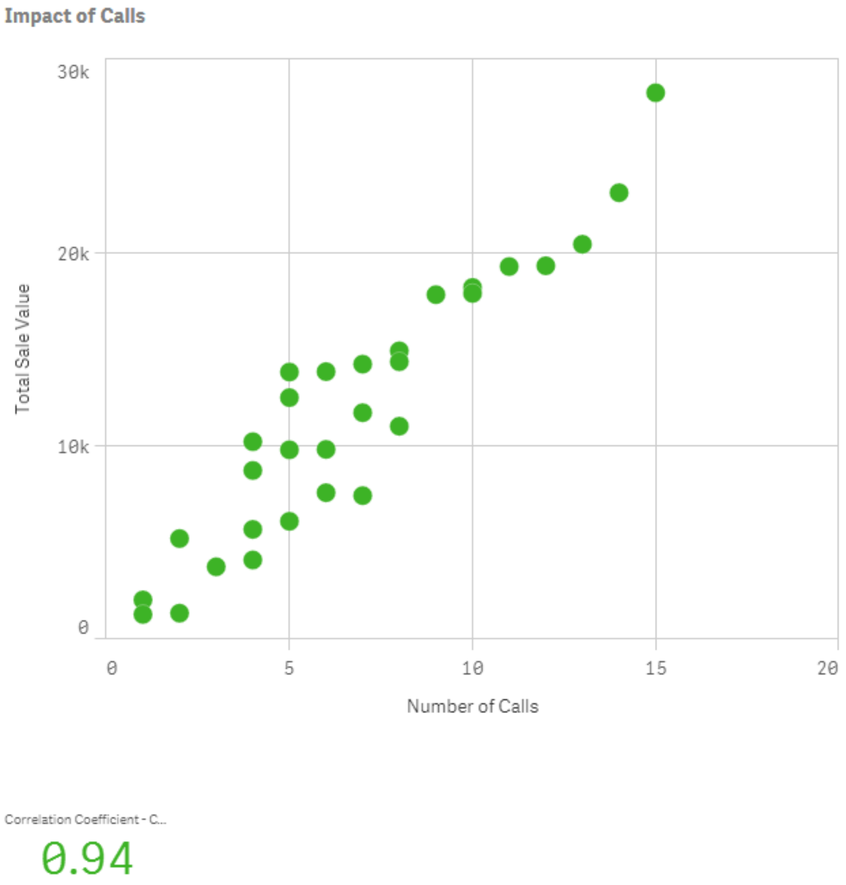
- Understanding data
- Understanding aggregation
- Understanding distribution

Fundamental analytics

- Signal & noise
- Correlation and causation
- Confidence intervals
- Analytical a/b testing
- Linear regression
- Hypothesis testing
- Design of experiment

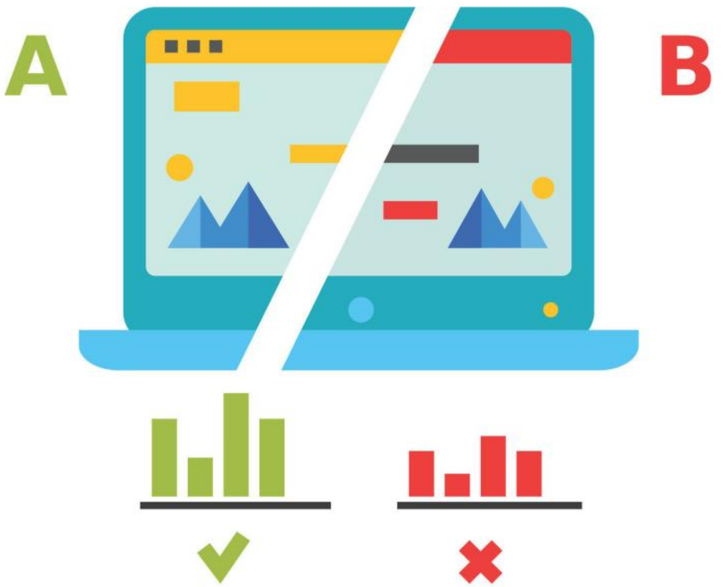
Chicken or the egg?

- Second most misused functions business have seen ☺
  - **Correlation** | not just scatter plot or trend line; correlation coefficient (r) <-1, +1> or 0; two variables behave similar
  - **Causation** | testing required in order to make conclusion
- **Use case |** What is the most productive time during the day for your plant workers? > **shift optimization, productivity planning**
  - Boss measures productivity > increase after daily stand-up at 1:15
  - One week they skip stand-up > productivity increases
  - TEST: 1 mth no stand-up & 2 weeks with 2 stand-ups > both increase > there was **correlation** but no **causation**
  - **Decision** > more smaller breaks to increase productivity & moral
- Other use cases: marketing campaigns, in-store promotions, new-product placements etc.



Ups, we changed something?

- Usual reaction on change ‘cos it’s not done right. > **a/b testing**
- **Use case |** New product launching, how to measure it?
  1. Define hypothesis what are you testing > **H0: our new Oreo chocolate will have at least 5% increase in 12 mth period compared to old Oreo**.
  2. Define size of control group > **it must be randomly selected**
  3. **Testing duration** > key to avoid biases, seasonality, football championships etc. (test one thing at a time)
  4. Analyze results > is the change **statistically significant**!?
  5. **Make decision**
- Other use cases: website changing, email conversion rate



Data fundamentals

- Understanding data
- Understanding aggregation
- Understanding distribution

Fundamental analytics

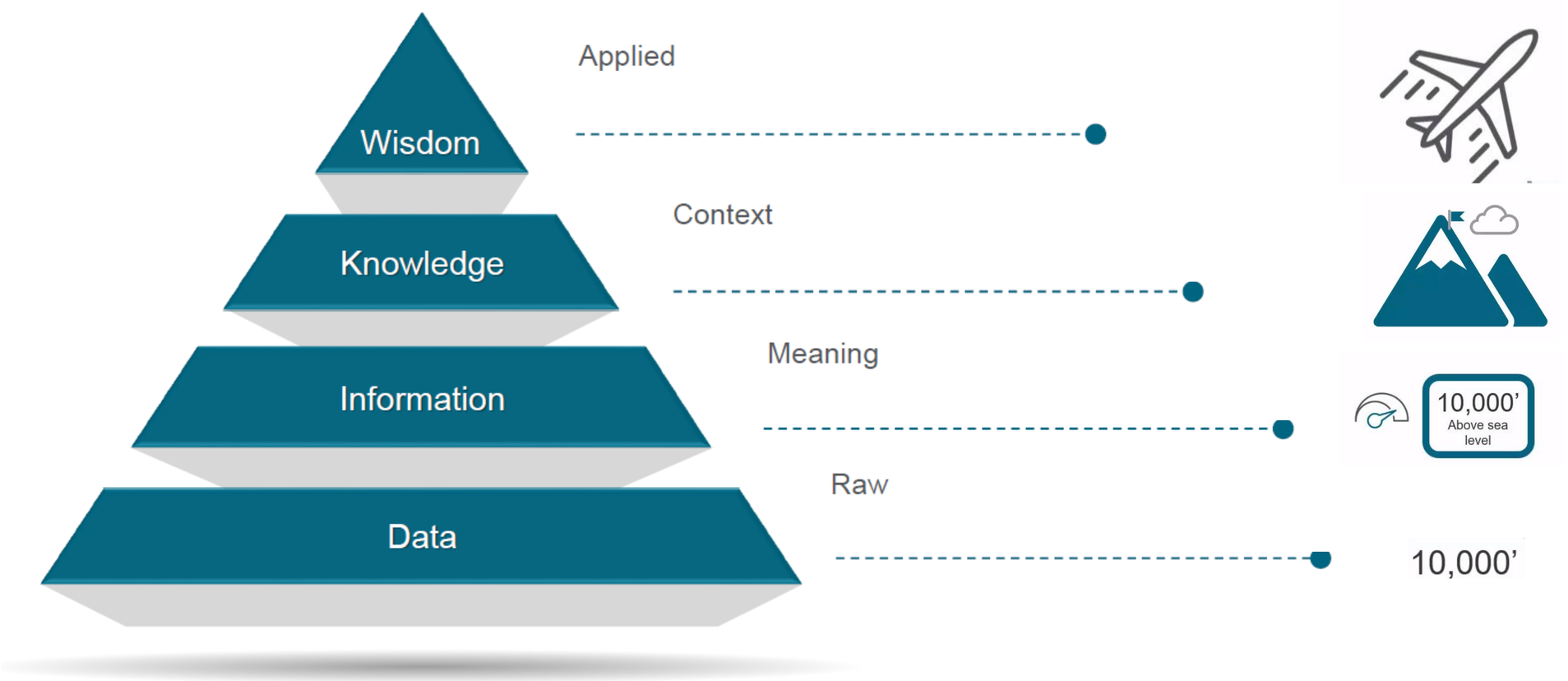
- Signal & noise
- Correlation and causation
- Confidence intervals
- Analytical a/b testing
- Linear regression
- Hypothesis testing
- Design of experiment

Data-informed decision making

- Introduction
- Framework
- Analytics techniques

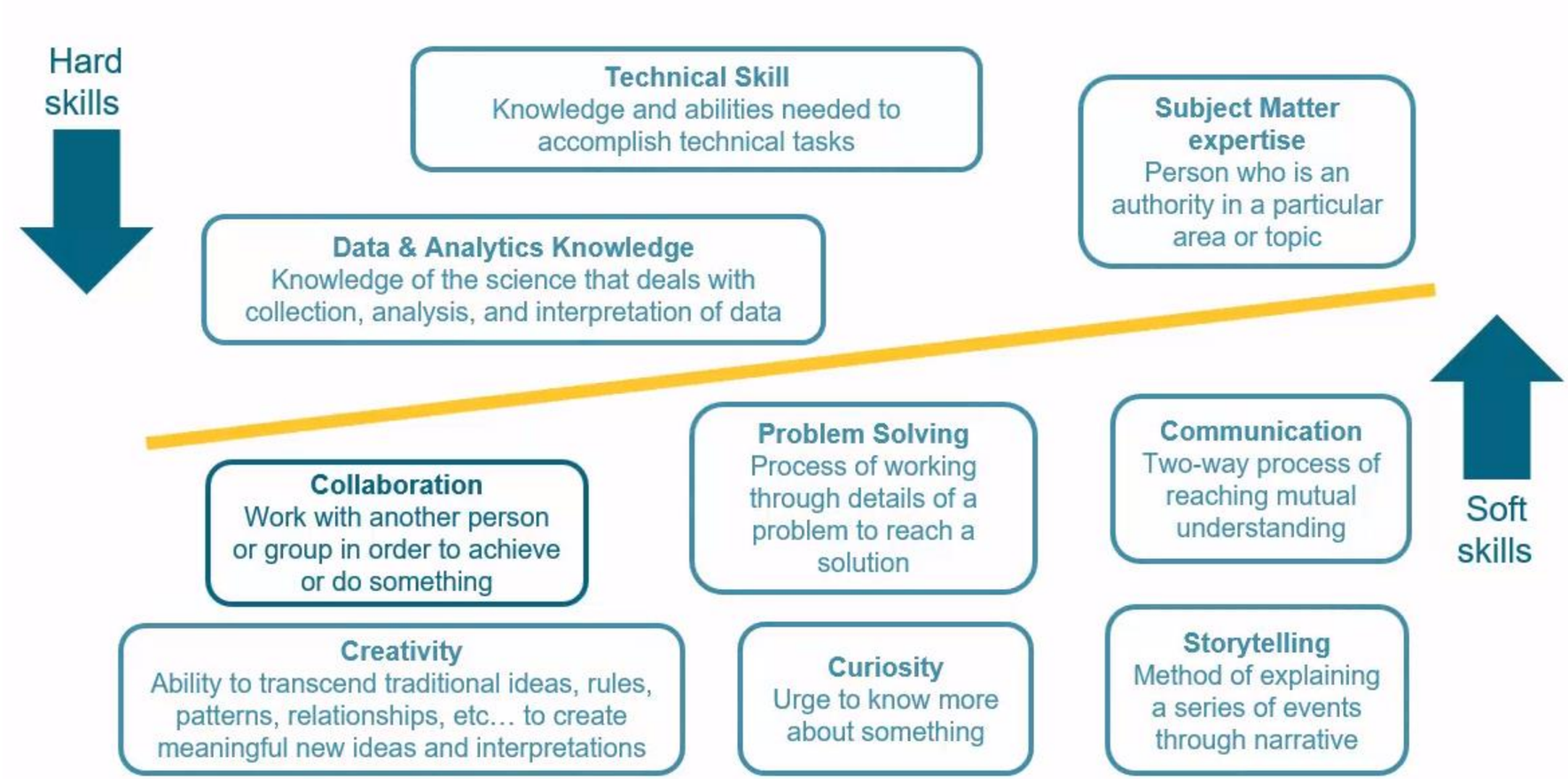
From data to wisdom (1)

- This process doesn't happen on it's own, we need to work on it
  > **hard & soft skills**

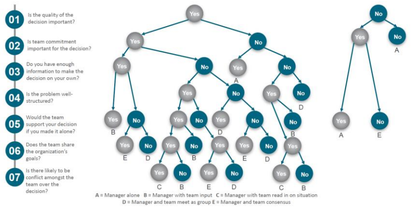


From data to wisdom (2)

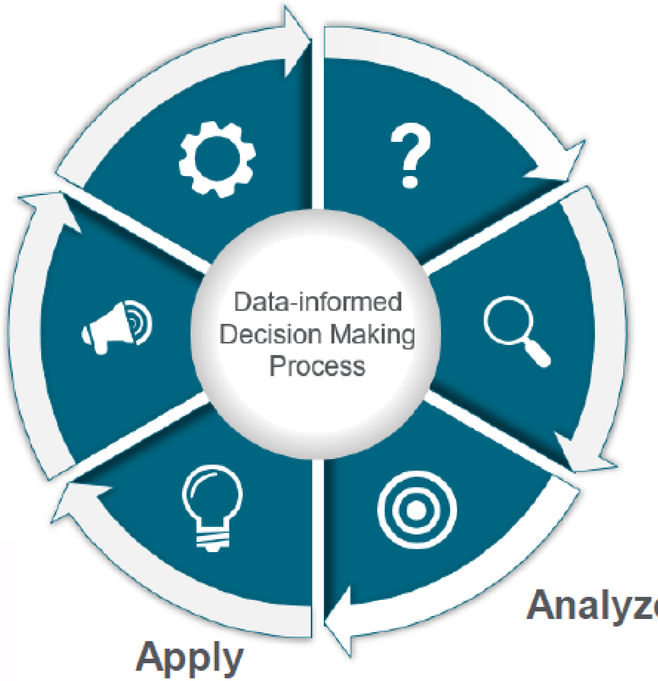
- This process doesn't happen on it's own, we need to work on it
  > **hard & soft skills**

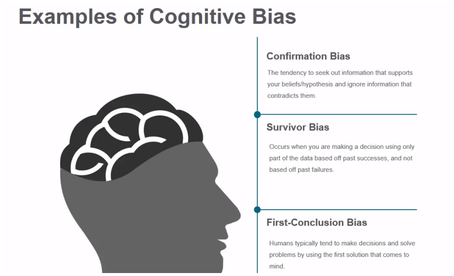


I hate frameworks, but

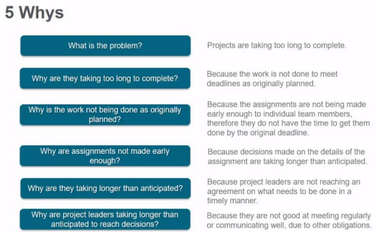


- Use the Vroom-Yetton-Jago Model to determine whether to make a decision alone or involve a group and to what extent the group should be involved



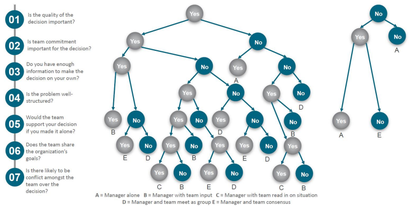


- Avoid cognitive biases: too much information, not enough meaning, and not enough time



- Use the 5 Whys technique to find the root cause of the problem

I hate frameworks, but



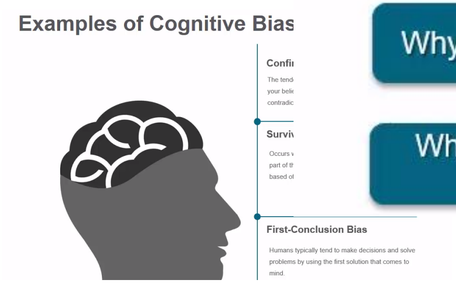
- Use the Vroom-Yetton-Jago Model to determine whe make a decision alone or involve a group and to wha the group should be involved

Anno

5 Whys

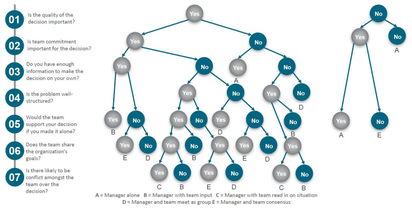
What is the problem?	Projects are taking too long to complete.
Why are they taking too long to complete?	Because the work is not done to meet deadlines as originally planned.
Why is the work not being done as originally planned?	Because the assignments are not being made early enough to individual team members, therefore they do not have the time to get them done by the original deadline.
Why are assignments not made early enough?	Because decisions made on the details of the assignment are taking longer than anticipated.
Why are they taking longer than anticipated?	Because project leaders are not reaching an agreement on what needs to be done in a timely manner.
Why are project leaders taking longer than anticipated to reach decisions?	Because they are not good at meeting regularly or communicating well, due to other obligations.

- Use the 5 Whys technique to find the root cause of the problem

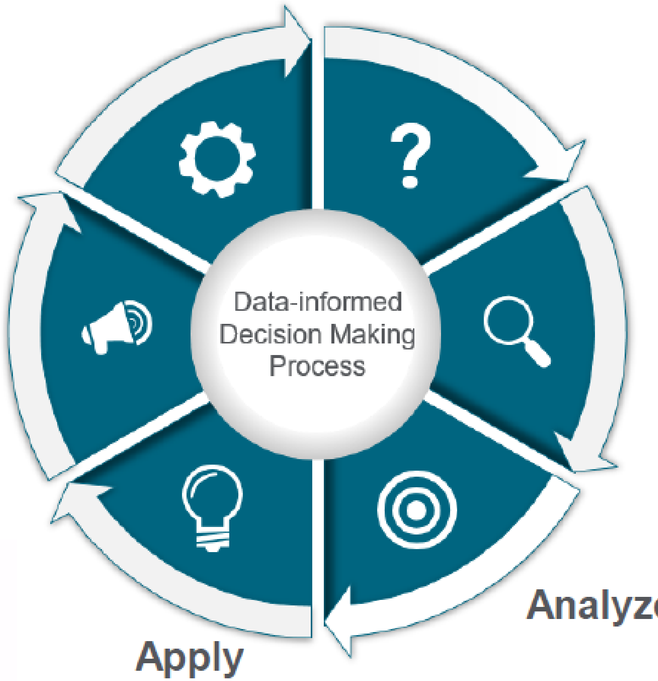


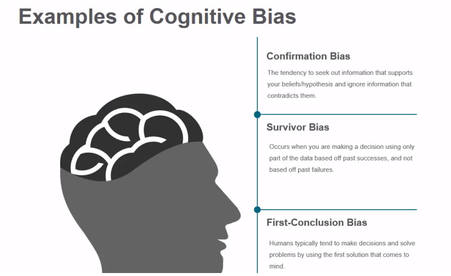
- Avoid cognitive biases: too much information, not enough meaning, and not enough time

I hate frameworks, but

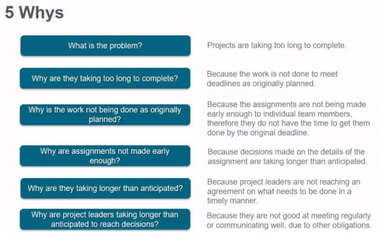


- Use the Vroom-Yetton-Jago Model to determine whether to make a decision alone or involve a group and to what extent the group should be involved



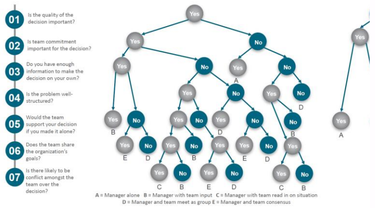


- Avoid cognitive biases: too much information, not enough meaning, and not enough time



- Use the 5 Whys technique to find the root cause of the problem

I hate frameworks, but



- Use the Vroom-Yetton-Jago
  make a decision alone or inv
  the group should be involved

Examples of Cognitive Bias

Confirmation Bias

The tendency to seek out information that supports your beliefs/hypothesis and ignore information that contradicts them.

Survivor Bias

Occurs when you are making a decision using only part of the data based off past successes, and not based off past failures.

First-Conclusion Bias

Humans typically tend to make decisions and solve problems by using the first solution that comes to mind.

Projects are taking too long to complete.

Because the work is not done to meet deadlines as originally planned.

Because the assignments are not being made early enough to individual team members, therefore they do not have the time to get them done by the original deadline.

Because decisions made on the details of the assignment are taking longer than anticipated.

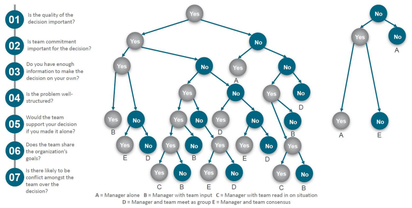
Because project leaders are not reaching an agreement on what needs to be done in a timely manner.

Because they are not good at meeting regularly or communicating well, due to other obligations.

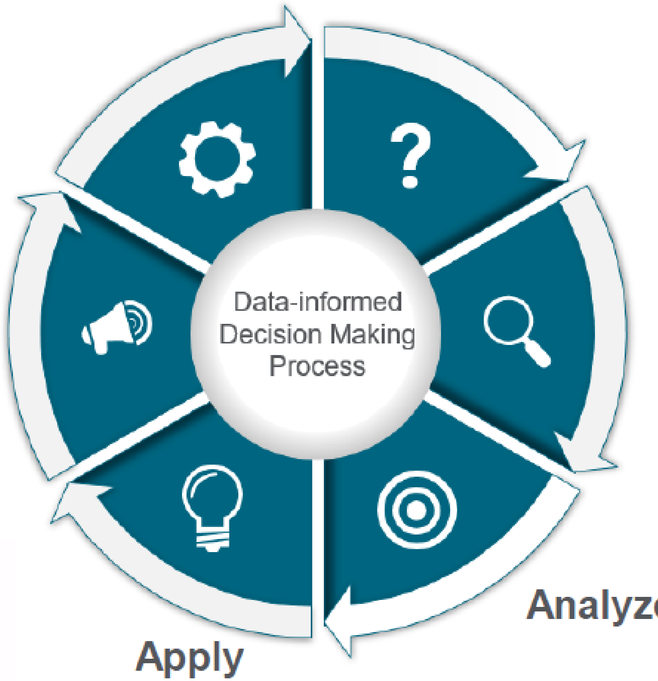
f the problem

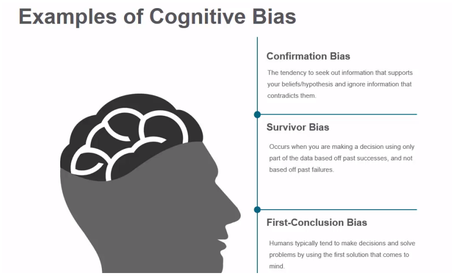
- Avoid cognitive biases: too much information, not enough meaning, and not enough time

I hate frameworks, but

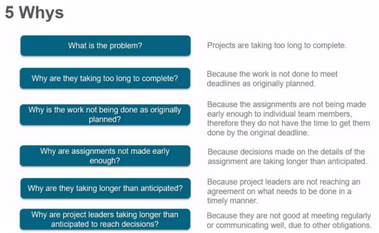


- Use the Vroom-Yetton-Jago Model to determine whether to make a decision alone or involve a group and to what extent the group should be involved





- Avoid cognitive biases: too much information, not enough meaning, and not enough time



- Use the 5 Whys technique to find the root cause of the problem

I hate frameworks, but

Assess

01 Is the quality of the decision important?

02 Is team commitment important for the decision?

03 Do you have enough information to make the decision on your own?

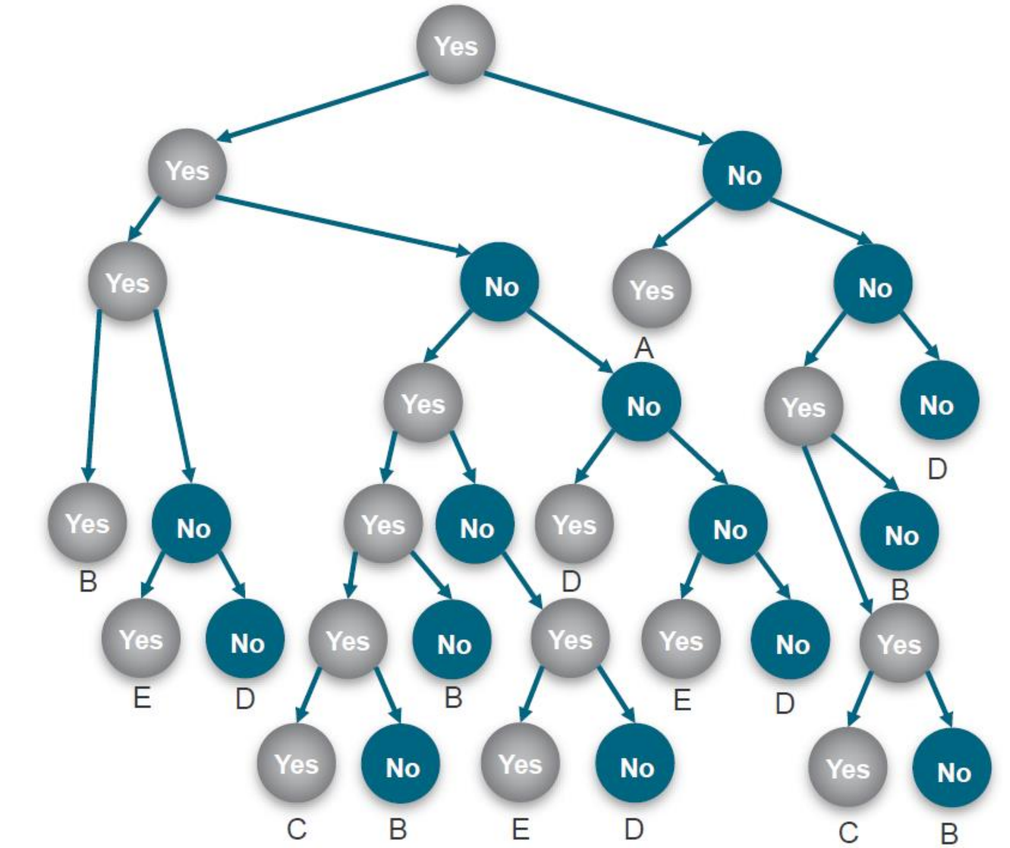
04 Is the problem well-structured?

05 Would the team support your decision if you made it alone?

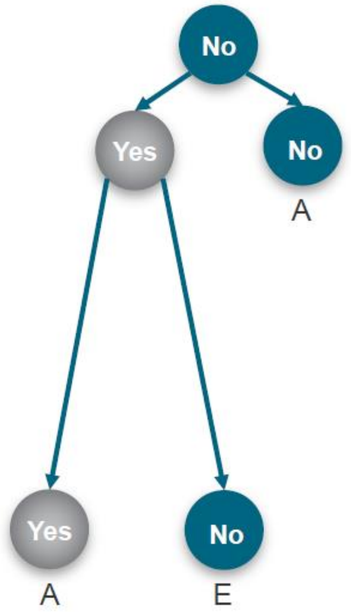
06 Does the team share the organization's goals?

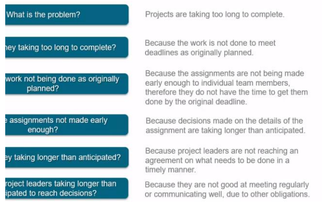
07 Is there likely to be conflict amongst the team over the decision?

Ask



A = Manager alone **B** = Manager with team input **C** = Manager with team read in on situation
D = Manager and team meet as group **E** = Manager and team consensus

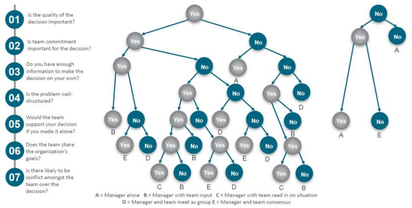




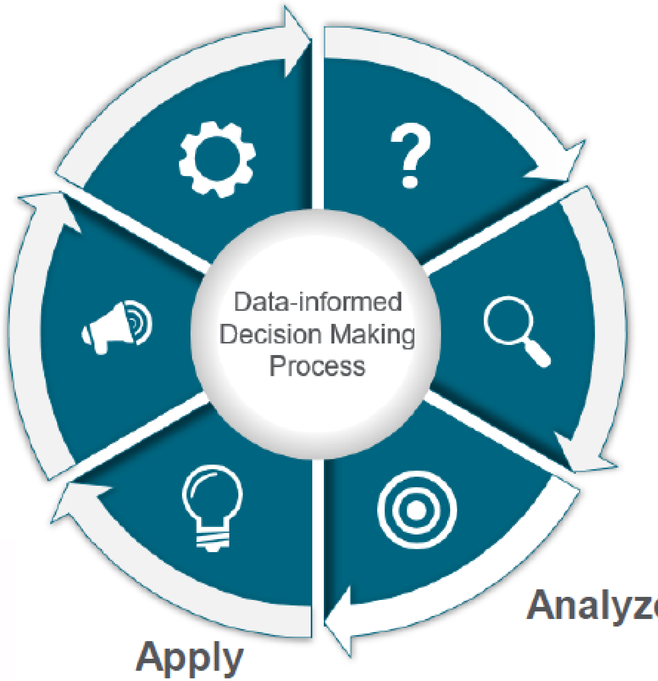
ot cause of the problem

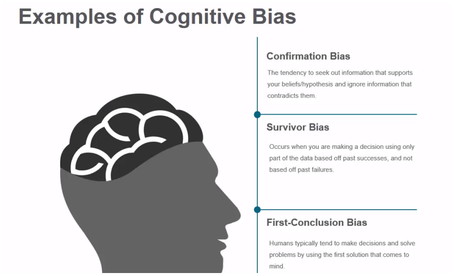
CONSIDER IT DONE

I hate frameworks, but

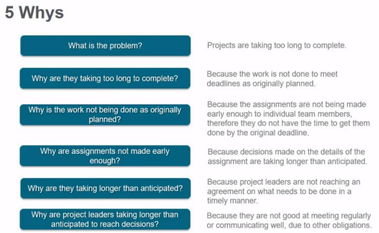


- Use the Vroom-Yetton-Jago Model to determine whether to make a decision alone or involve a group and to what extent the group should be involved



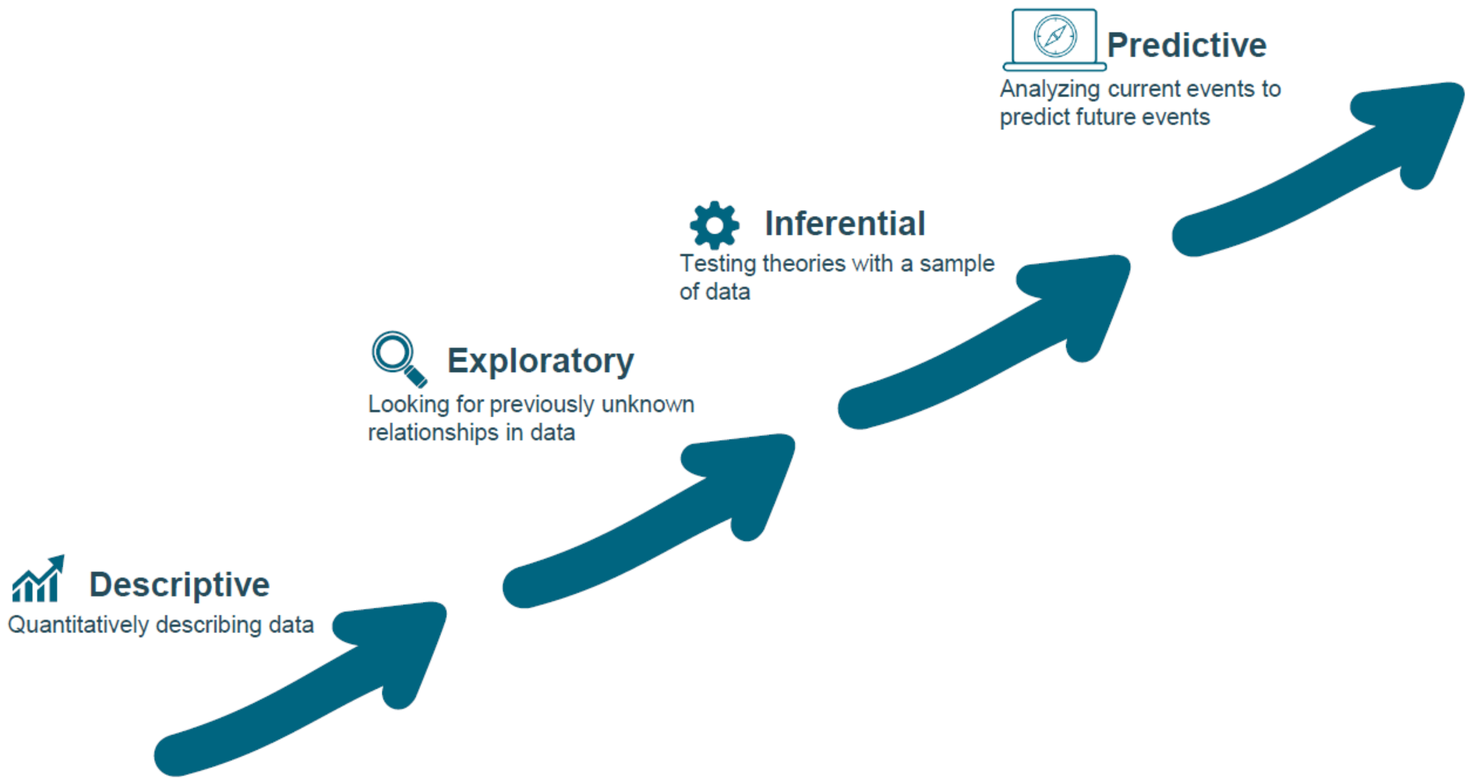


- Avoid cognitive biases: too much information, not enough meaning, and not enough time



- Use the 5 Whys technique to find the root cause of the problem

It all matters



- Predictive
  - Breakeven, sentiment analysis, NLP, machine learning
  - **Food manufacturing |** What is the optimal sequence of workorders in order to maximize production in one day?
- Inferential
  - Hypothesis, significance, sample/population
  - **Automotive |** What is the probability that new car model will cannibalize current one and in which amount?
- Exploratory
  - Correlation, regression etc.
  - **Services & consulting |** Why are we behind the plan for 3 months already?
- Descriptive
  - Sum, Average, Mean, Percentage % etc.
  - **Retail |** What was my result last month?

CONSIDER IT DONE