Data Science and Project Management

North West Project Data Analytics Meetup

Aims

1. How to get started in data science
2. Why data science in project management makes sense
3. How to undertake a program of intelligent automation and AI in project management.
4. The caveats and gotchas of implementation of intelligent automation and AI
5. The spoils for companies who undertake an AI and intelligent automation transformation

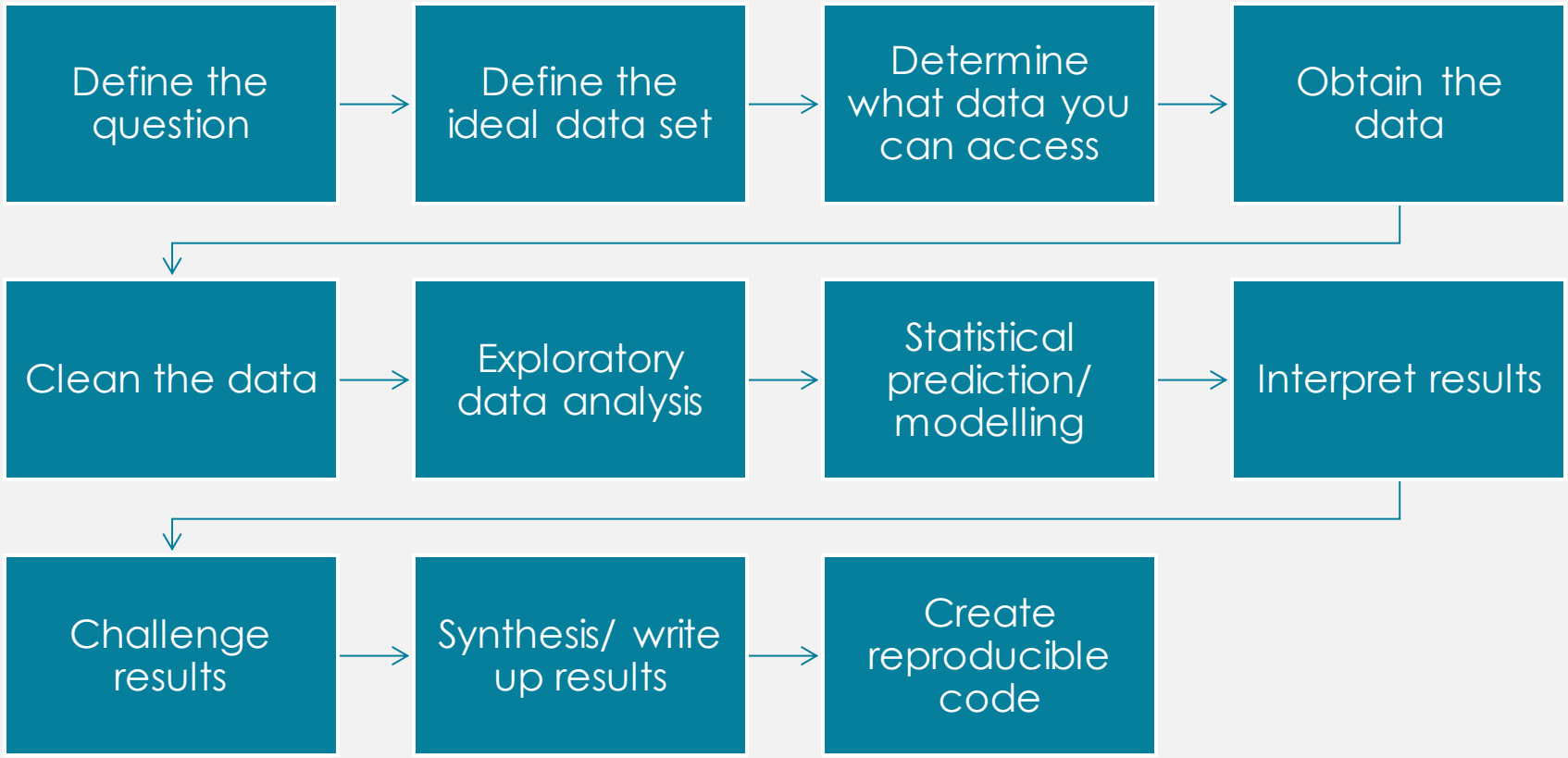
1. Getting started in data science

Or at least how I did it.

I) What is data science?

Using the data assets of a business to help the business achieve its strategic aims.

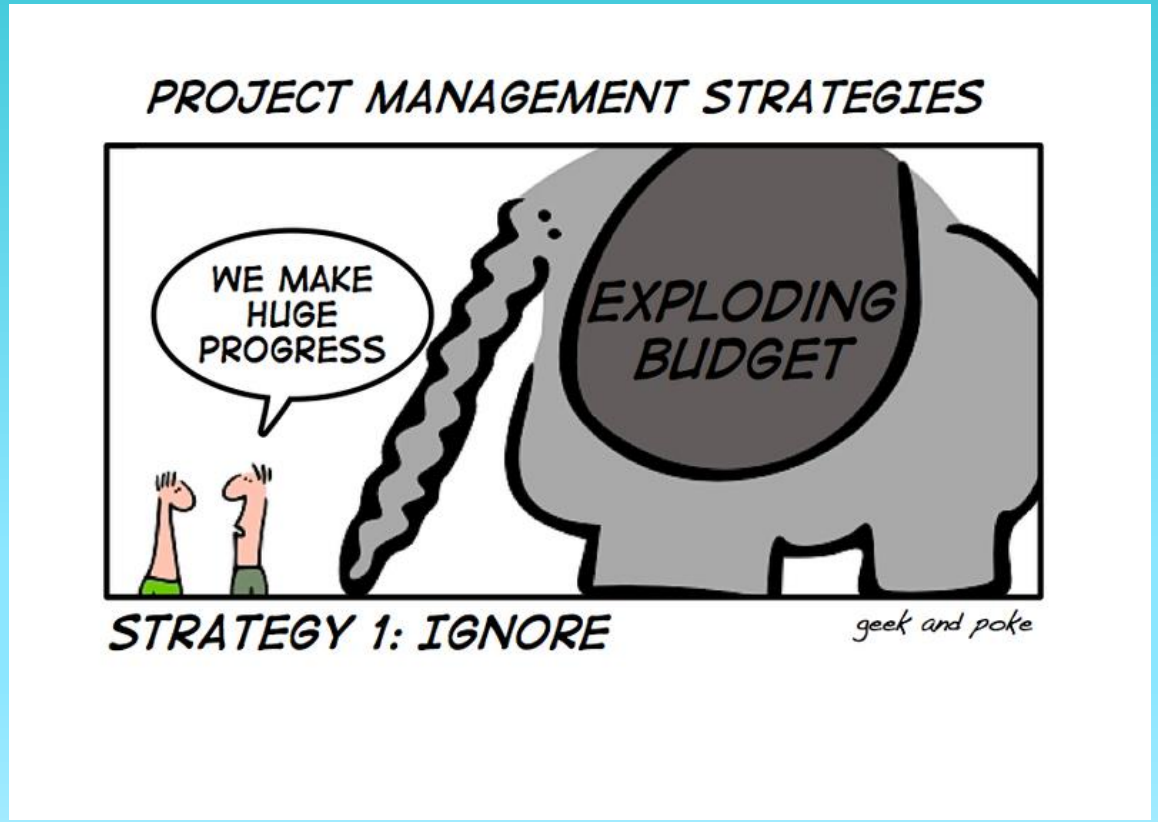
II) Steps in the data science process



III) 80/20 data science skills

Command line	Git version control	SQL and database concepts
You can read data into R/ Python	You can manipulate data in R/ Python	You can plot data in R/ Python
You can fit basic models in R/ Python	You can document your results and can reproducible code in R/ Python	You can present results

2. Data science in project management

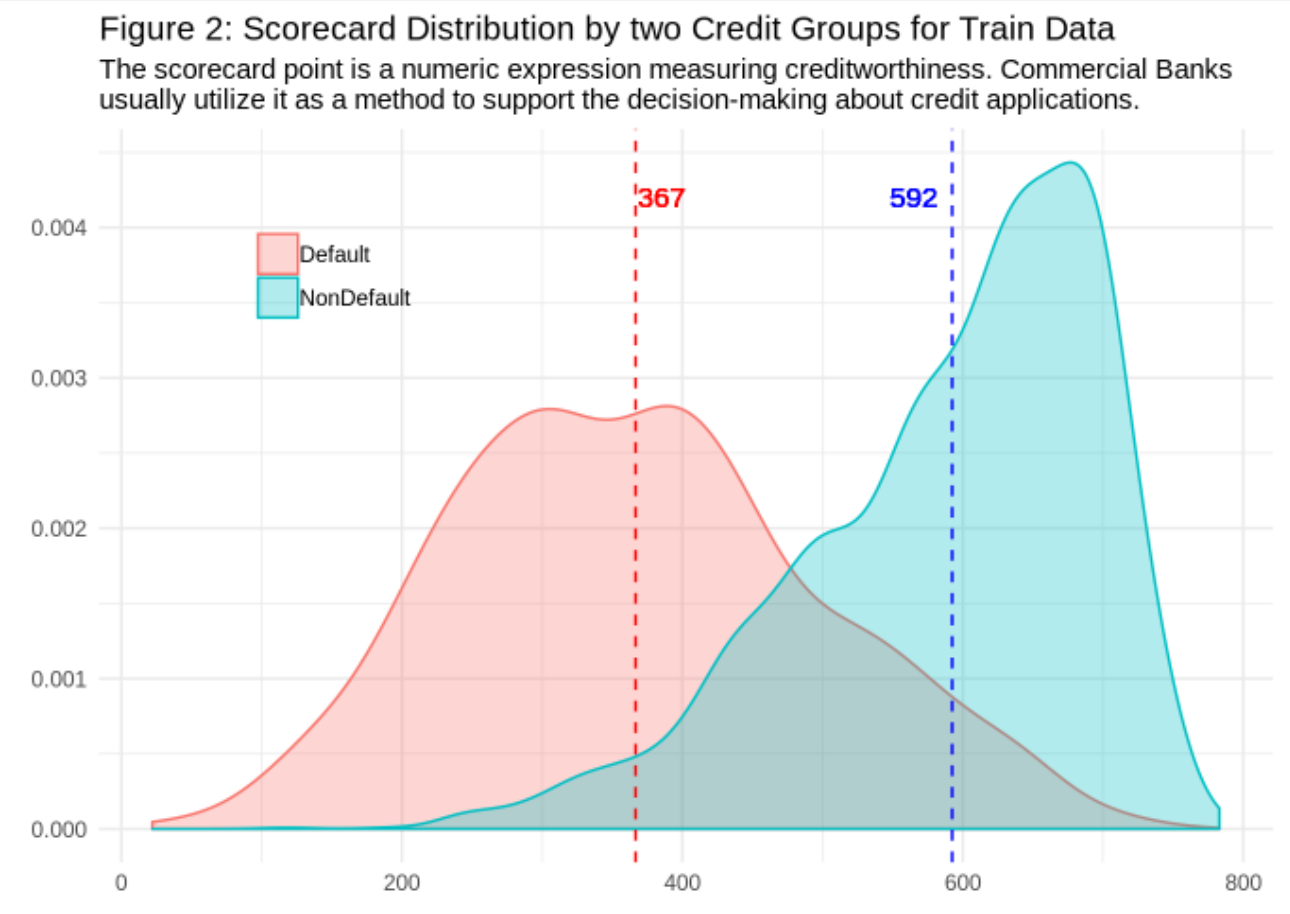


I) The biggest industries in the UK

1. Finance and Banking
2. IT
3. Construction
4. Oil and Gas
5. Government
6. Healthcare
7. Manufacturing
8. Wholesale and Retail
9. Transportation and Logistics
10. Education

II) You can't model projects

1. Relationship lending in banking
    1. Bias in credit decisioning
    2. Automated credit scoring uses past loan performance on application information
2. Human judgement still needed for decisions
    1. https://rpubs.com/chidungkt/442168
    2. Less than 367 is an automated rejection
    3. More than 592 is an automated acceptance
    4. Between 367 and 592 requires more investigation or documentation, this is the place for human judgement.



II) You can't model projects

Characteristic	Attribute	Scorecard Points
AGE	<22	100
AGE	22<=AGE<26	120
AGE	26<=AGE<29	185
AGE	29<=AGE<32	200
AGE	32<=AGE<37	210
AGE	37<=AGE<42	225
AGE	>=42	250
HOME	OWN	225
HOME	RENT	110
INCOME	<10000	120
INCOME	10000<=INCOME<17000	140
INCOME	17000<=INCOME<28000	160
INCOME	28000<=INCOME<35000	200
INCOME	35000<=INCOME<42000	225
INCOME	42000<=INCOME<58000	230
INCOME	>=58000	260

Let cutoff=600

So, a new customer applies for credit.....

AGE	35	210 points
INCOME	$38K	225 points
HOME	OWN	225 points
Total		660 points
Decision:		GRANT CREDIT

Note: A scorecard is scaled with the **Odds**, **Scorecard Points** and **Points to Double the Odds** properties.

1. The modelling exercise gives insight into the variables related to credit worthiness
    1. Younger people are less credit worthy
    2. Renters are higher risk
    3. People with lower incomes are riskier

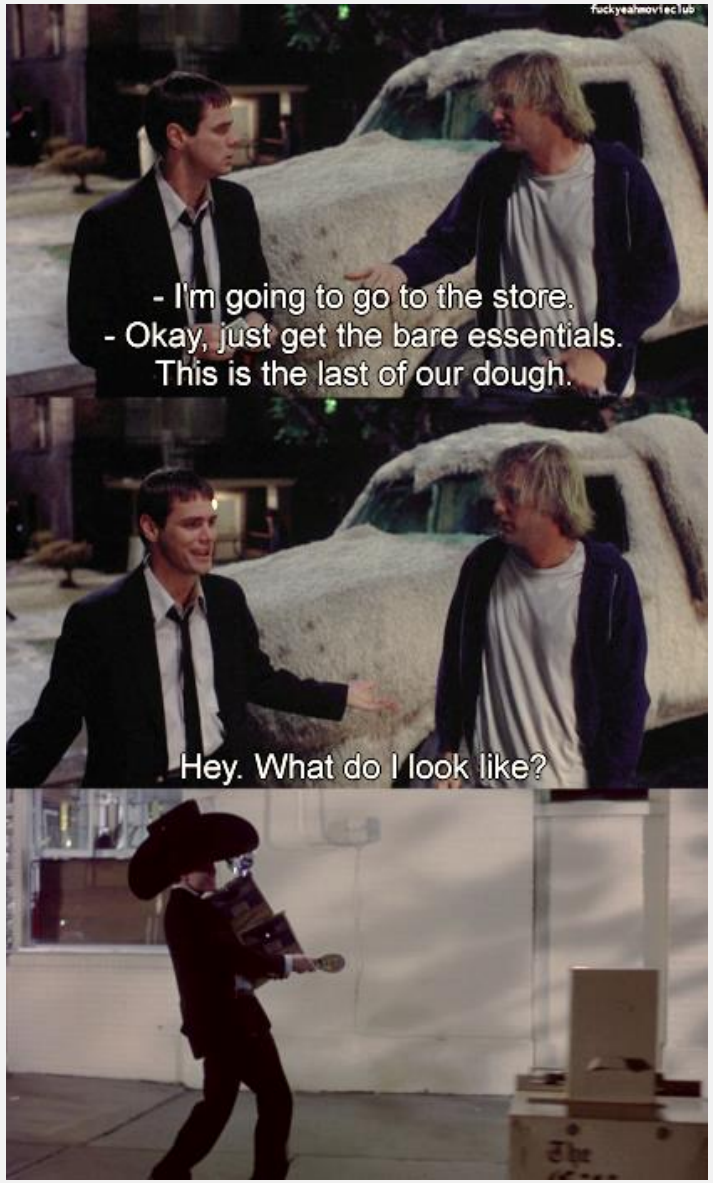
https://medium.com/@yanhuiliu104/credit-scoring-scorecard-development-process-8554c3492b2b

III) Each project is unique

1. Discretionary spending is the lead indicator of default in lending
    1. Across socioeconomic factors
    2. Across locations
    3. Across any variable I could think of

2. There are tell tale signs of credit default, we may also see signs of project overrun across projects
    1. Increased demand for new credit
    2. Decrease in net cash
    3. A credit card default will precede a home loan default

3. Pool of risks in insurance, reinsurance
    1. Each company may insure a couple Ferraris, but across the country there are enough Ferraris to model insurance prices.
    2. Maybe a Porsche, Ferrari and Lamborghini can be pooled together, similar characteristics?



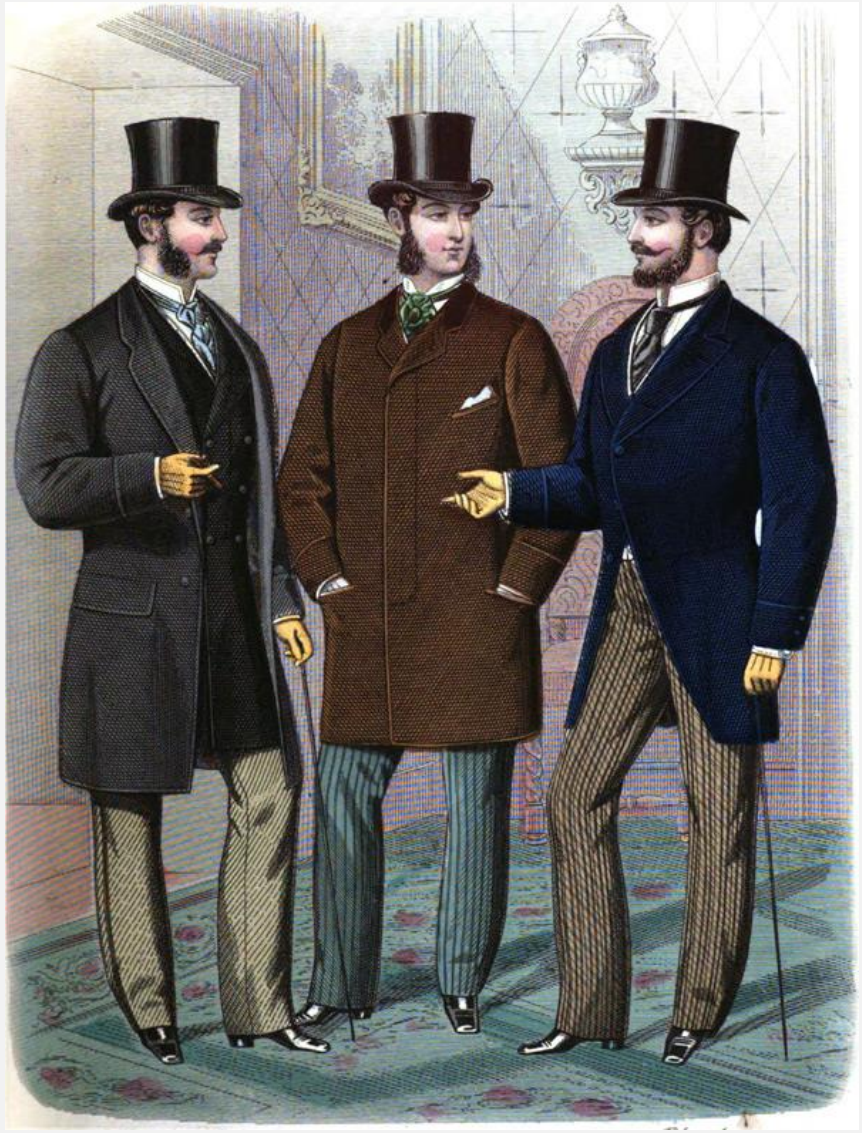
IV) Your data is an asset of your business

1. The modern credit bureau was started by a group of English tailors who shared information on customers who had ripped them off.

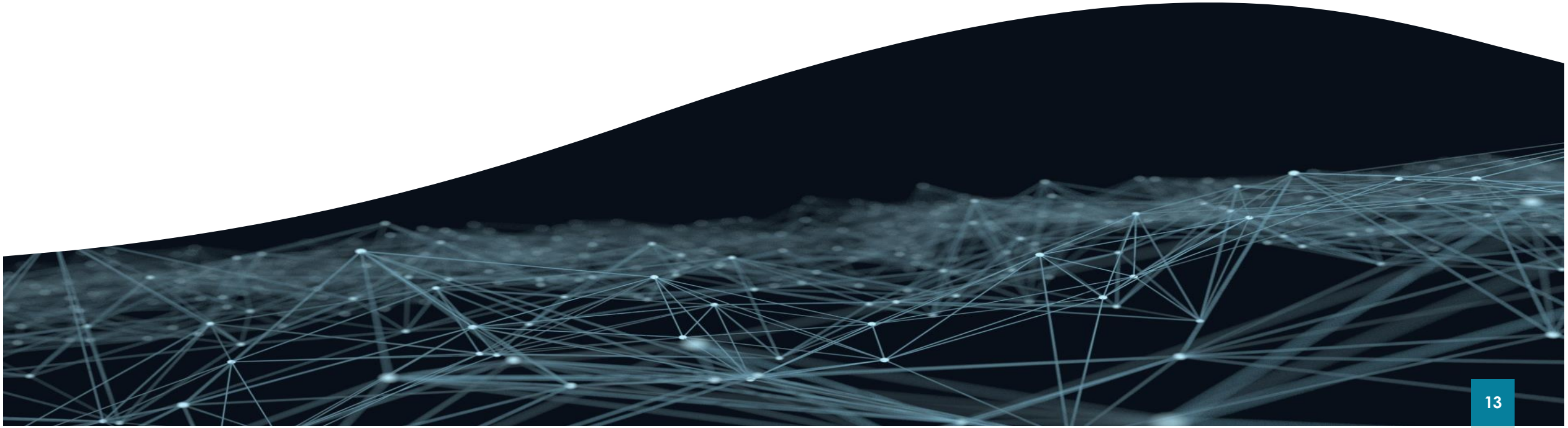
2. Likewise collecting the information you have on past projects can be immense value in the future.

3. Algorithms and tools are cheap or free, education and training is available. Your data is what is valuable.

https://www.historicalemporium.com/mens-late-victorian-clothing.php

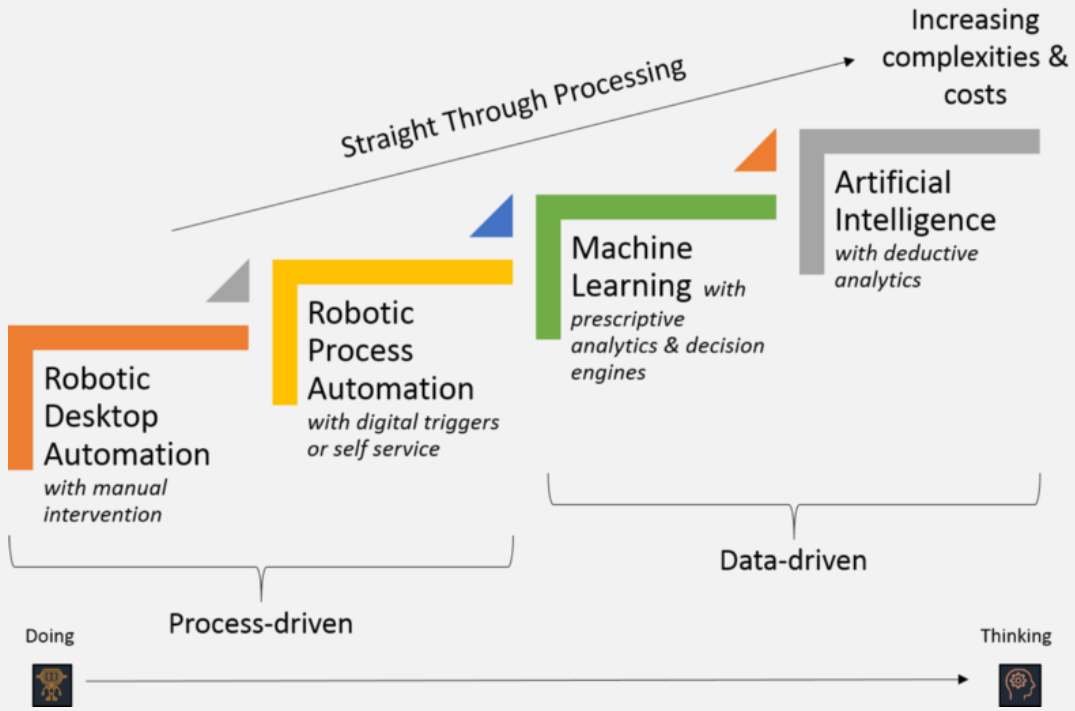


3. AI implementation in project management



I) A roadmap for intelligent automation and AI in project management

1. Start by automating Excel
2. Build dashboards with tools like Power BI
3. Automate entire processes with RPA
4. Use data to answer business questions with data science projects.
5. Take insights from data science to actions with A/B testing
6. Production machine learning for real time decisions



II) Caveats and gotchas

1. Cascade the corporate strategy to the Data Science Team. A misaligned corporate strategy is toxic.

2. Eat an elephant one bite at a time in 2 week sprints with a clear deliverable at the end.

3. Recruit the right people at the right time.

4. Beware the “vapor ware” vendors

III) The implications for project management

1. Better data capture and quality allowing the organization to mov e down the path of intelligent automation.

2. Project managers informed by data rather than wasting time wrangling data

3. More projects completed on time and at reduced risk, including reputational risk.

4. Understanding of pain points in projects

Conclusion

1. How to get started in data science
2. Why data science in project management makes sense
3. How to undertake a program of intelligent automation and AI in project management.
4. The caveats and gotchas of implementation of intelligent automation and AI
5. The spoils for companies who undertake an AI and intelligent automation transformation