Pengantar Business & Data Analytics
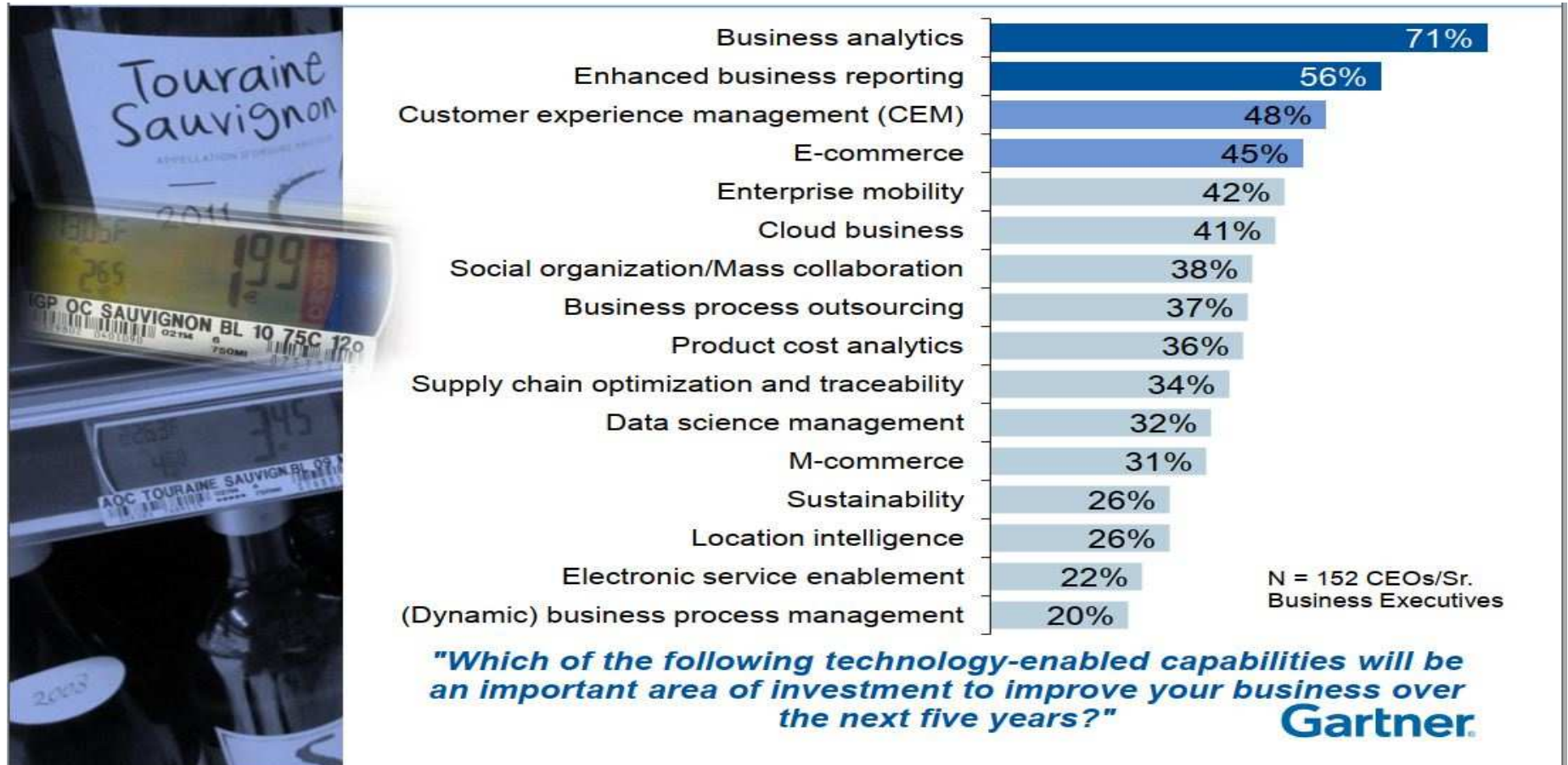
ASTRA HONDA MOTOR — FINANCE DEPARTMENT
JAKARTA, FEBRUARI 2019

Natanael Peranginangin
E-mail:
natanael.peranginangin@gmail.com

Where to Invest to Improve Business?

Business analytics — 71%
Enhanced business reporting — 56%
Customer experience management (CEM) — 48%
E-commerce — 45%
Enterprise mobility — 42%
Cloud business — 41%
Social organization/Mass collaboration — 38%
Business process outsourcing — 37%
Product cost analytics — 36%
Supply chain optimization and traceability — 34%
Data science management — 32%
M-commerce — 31%
Sustainability — 26%
Location intelligence — 26%
Electronic service enablement — 22%
(Dynamic) business process management — 20%

N = 152 CEOs/Sr. Business Executives

"Which of the following technology-enabled capabilities will be an important area of investment to improve your business over the next five years?"
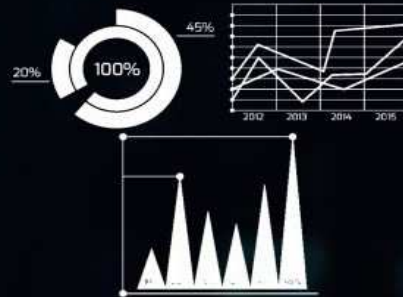
Gartner.

Traditional Data Analyst

A **Data Analyst** takes data and uses it to help companies make better **business decisions.**
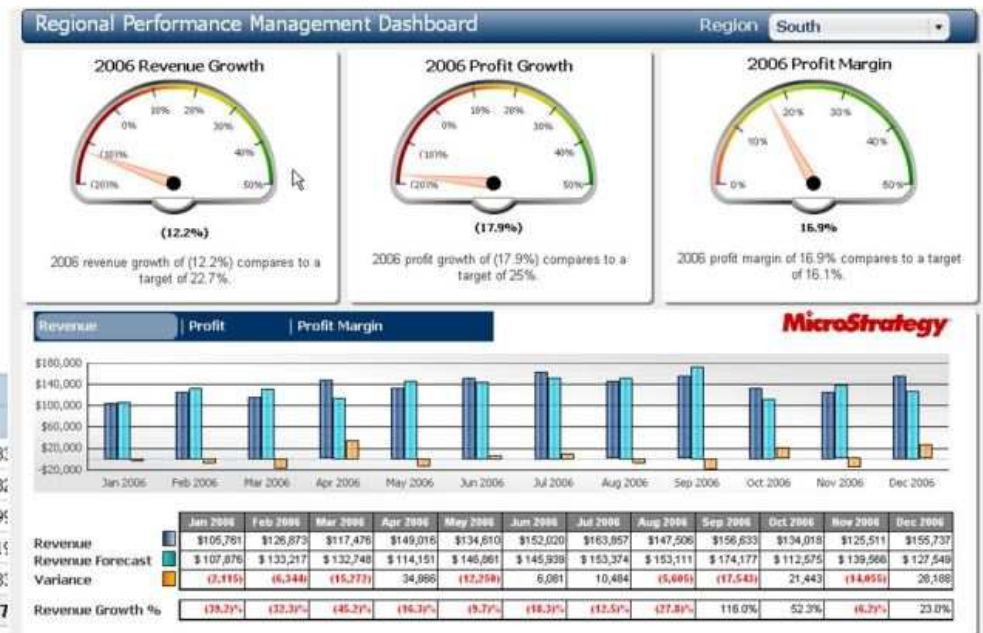
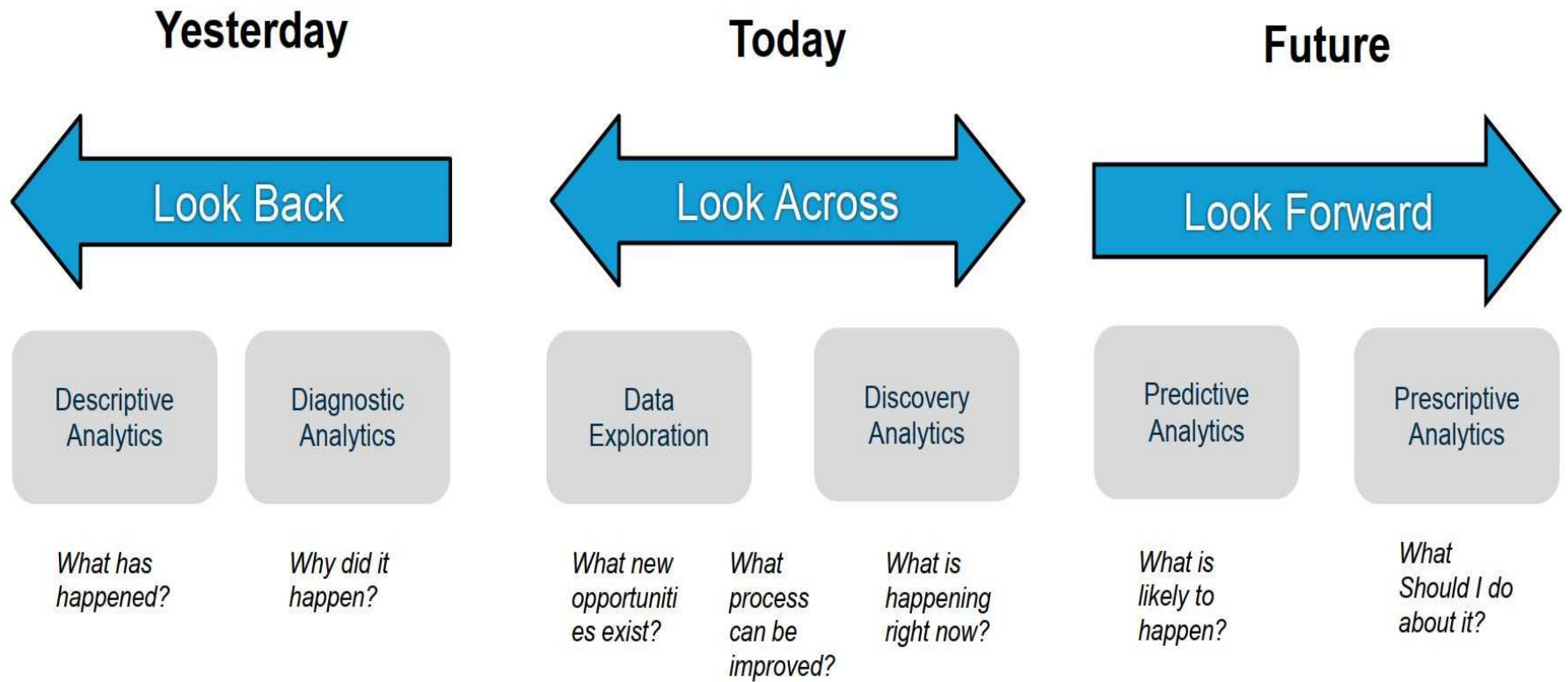Collect Data | Analyse Data | Create Reports
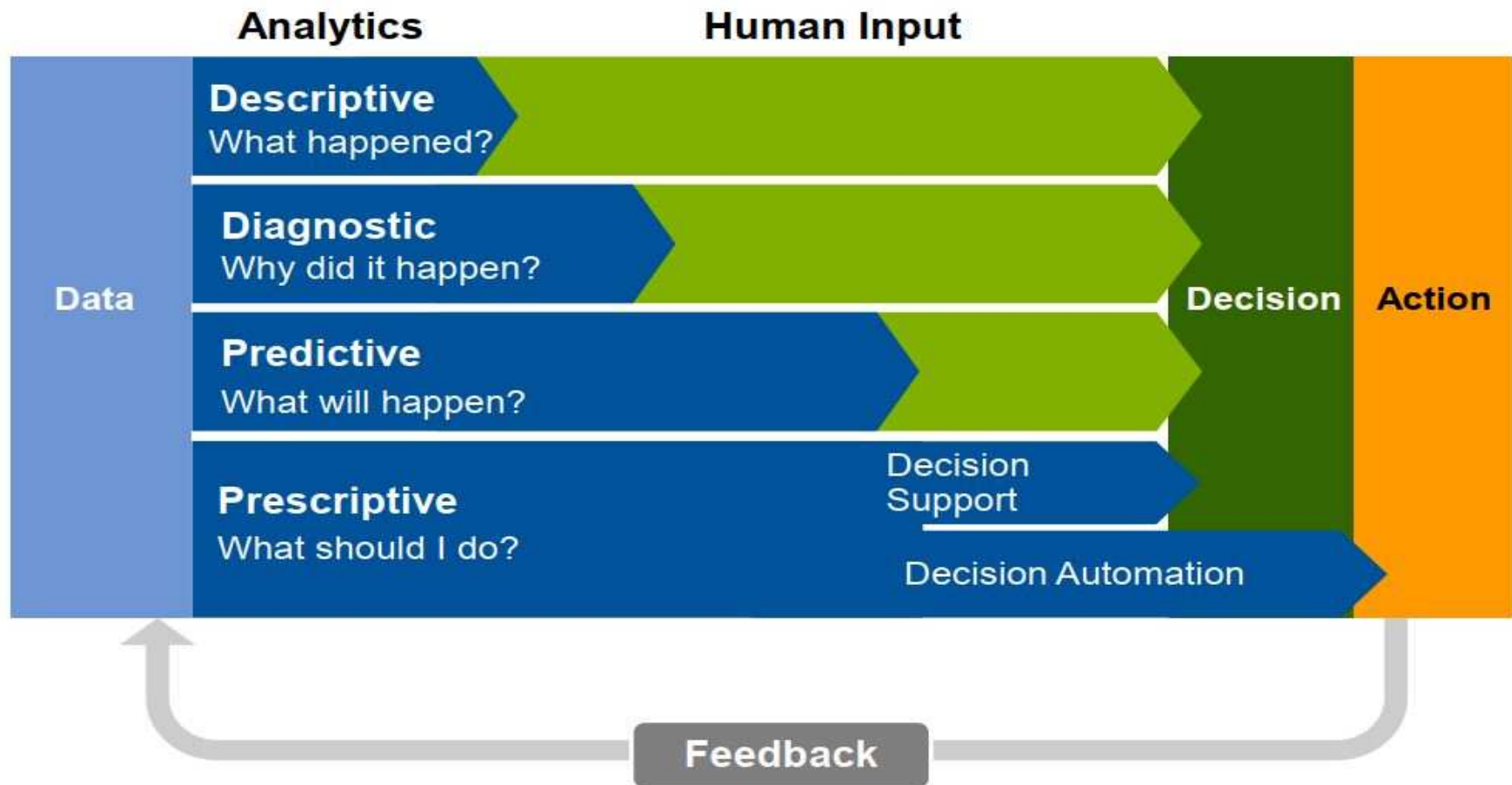
Report Example

Expanding Role of Data Analyst

Yesterday

Today

Future

Look Back

Look Across

Look Forward

Descriptive Analytics	Diagnostic Analytics	Data Exploration	Discovery Analytics	Predictive Analytics	Prescriptive Analytics
What has happened?	*Why did it happen?*	*What new opportunities exist?*	*What process can be improved?* *What is happening right now?*	*What is likely to happen?*	*What Should I do about it?*

Towards Automated Decision

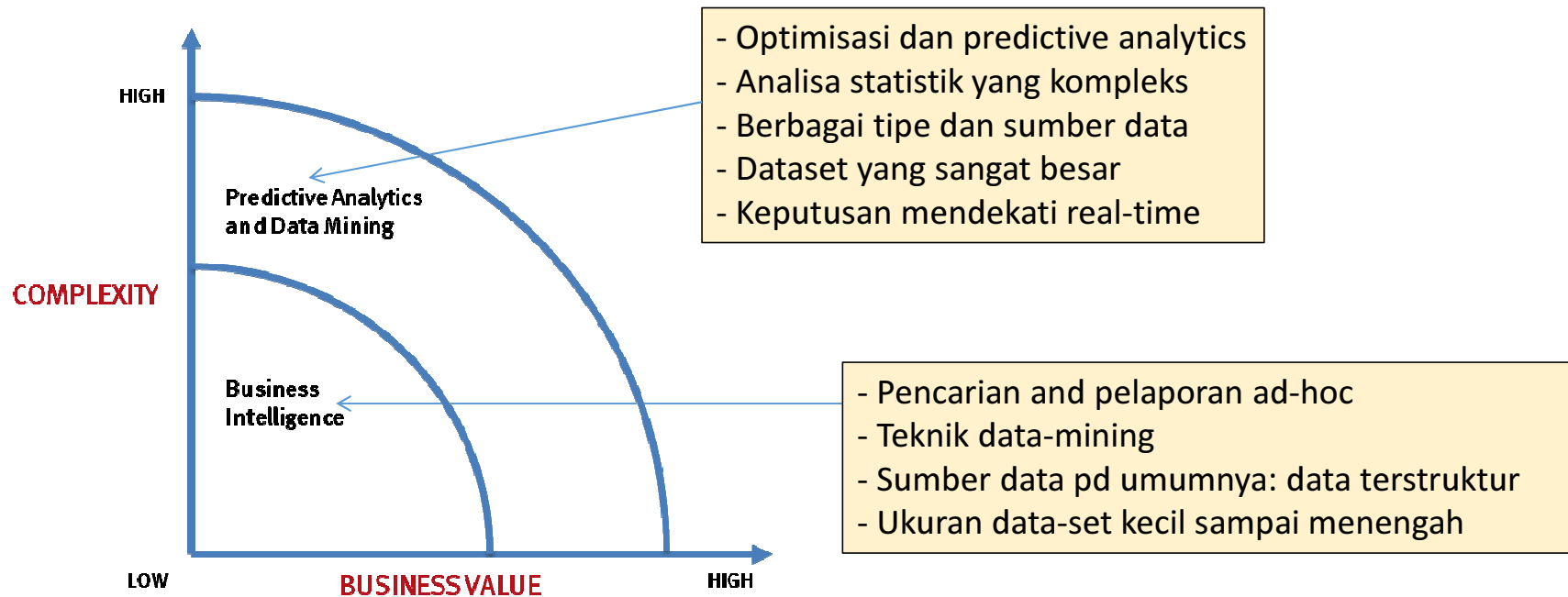# Intelligence vs. Analytics

Traditional BI
- Standard reports and dashboards
- Ad hoc reports – Current performance
- Query Drill down
- Cube analysis – Slice and dice
- Alerts

Vs.

Business Analytics
- Statistical Analysis
- Forecasting
- Predictive modeling
- Optimization

Intelligence vs Analytics

COMPLEXITY (vertical axis): LOW to HIGH
BUSINESS VALUE (horizontal axis): LOW to HIGH

Predictive Analytics and Data Mining

Business Intelligence

- Optimisasi dan predictive analytics
- Analisa statistik yang kompleks
- Berbagai tipe dan sumber data
- Dataset yang sangat besar
- Keputusan mendekati real-time

- Pencarian and pelaporan ad-hoc
- Teknik data-mining
- Sumber data pd umumnya: data terstruktur
- Ukuran data-set kecil sampai menengah

8

Data Landscape Today

BIG DATA?

VOLUME
Large amounts of data.
VELOCITY
Needs to be analyzed quickly.
VARIETY
Different types of structured and unstructured data.

Complexity

Big Data

Speed   Volume

Big Data = Transactions + Interactions + Observations

BIG DATA

Petabytes
Sensors / RFID / Devices
Mobile Web
User Click Stream
Sentiment
User Generated Content
Social Interactions & Feeds
Spatial & GPS Coordinates

Terabytes
Web logs
Offer history
WEB
A/B testing
Dynamic Pricing
External Demographics
Business Data Feeds

Gigabytes
CRM
Segmentation
Offer details
Affiliate Networks
Search Marketing
Behavioral Targeting
HD Video, Audio, Images
Speech to Text

Megabytes
ERP
Purchase detail
Purchase record
Payment record
Customer Touches
Support Contacts
Dynamic Funnels
Product/Service Logs
SMS/MMS

Increasing Data Variety and Complexity

Source: Contents of above graphic created in partnership with Teradata, Inc.

Economics of Data Analytics
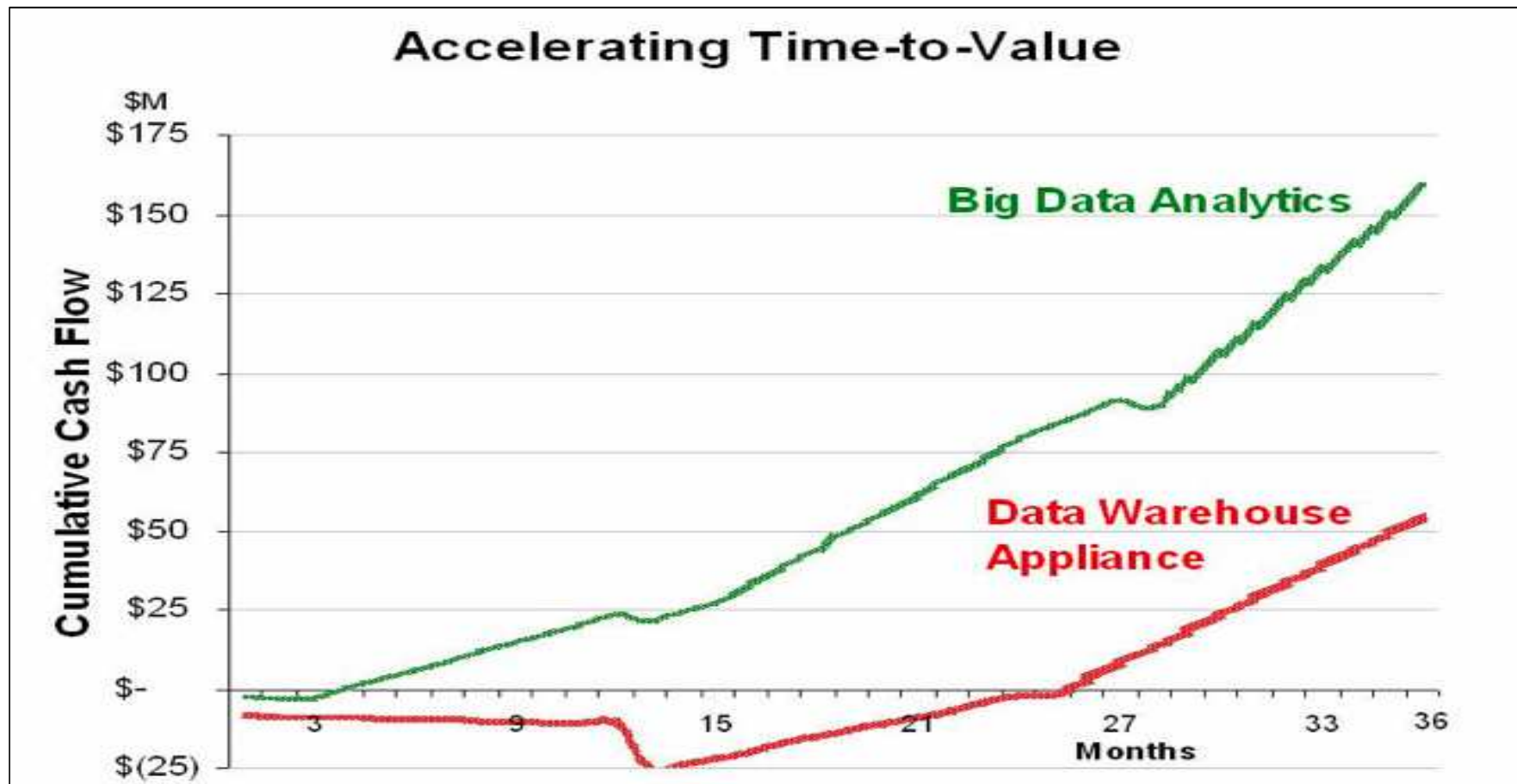
Beneficiary
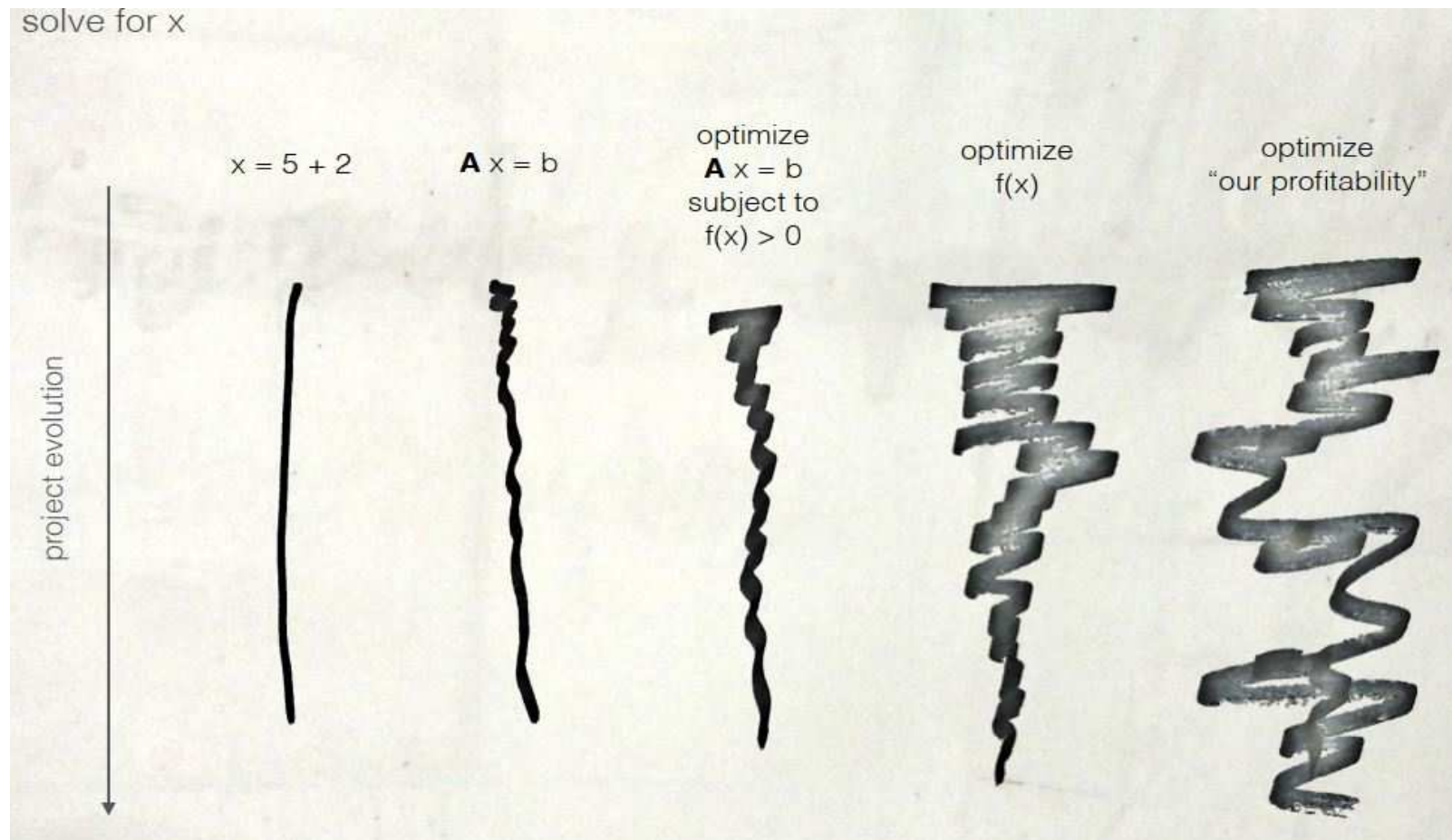
1) Analytics for Humans

- Another **human is the final decision maker** and consumer of the analysis

- Must be **comfortable coming to higher-level conclusions – the "why" and "how"**

- Telling a story from the data

Beneficiary

2) Analytics for Machines

- The **final decision maker** and consumer of the analysis **is a computer**
- Creating computer **algorithms and models**
- **Their digital models are established and then act on their own**
    - automatically trade in the stock market
    - Decide ads to display for online content/advertising targeting, or
    - Personalized product recommendations
- must have remarkably **strong mathematical, computational, and statistical skills and data modelling**
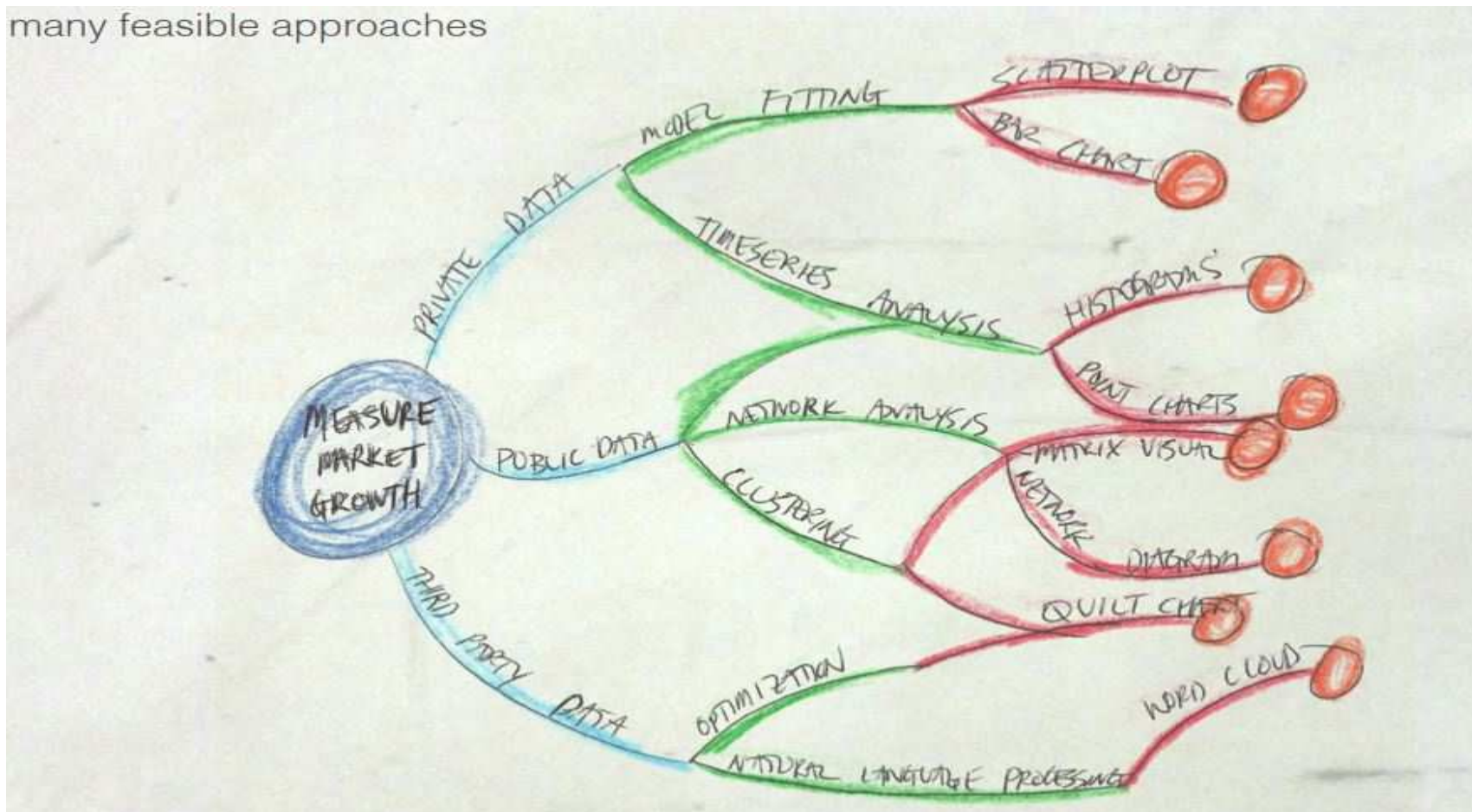- So that systems can make quality predictions quickly.

Typical Problem to Solve

solve for x

project evolution

| $x = 5 + 2$ | $\mathbf{A}x = b$ | optimize $\mathbf{A}x = b$ subject to $f(x) > 0$ | optimize $f(x)$ | optimize "our profitability" |

# Problem Type

Type of Approach

many feasible approaches

Thinking Guidelines

- **Context:** What are you trying to achieve? Who is invested in the project's results? Are there any larger goals or deadlines that can help prioritize the project?

- **Need:** What specific needs could be addressed by intelligently using data? What will this project accomplish that was impossible before?

- **Vision:** What will meeting the need with data look like? Is it possible to mock up the final result? What is the logic behind the solution?

- **Outcome:** How and by whom will the result be used and integrated into the company? How will the success of the project be measured?
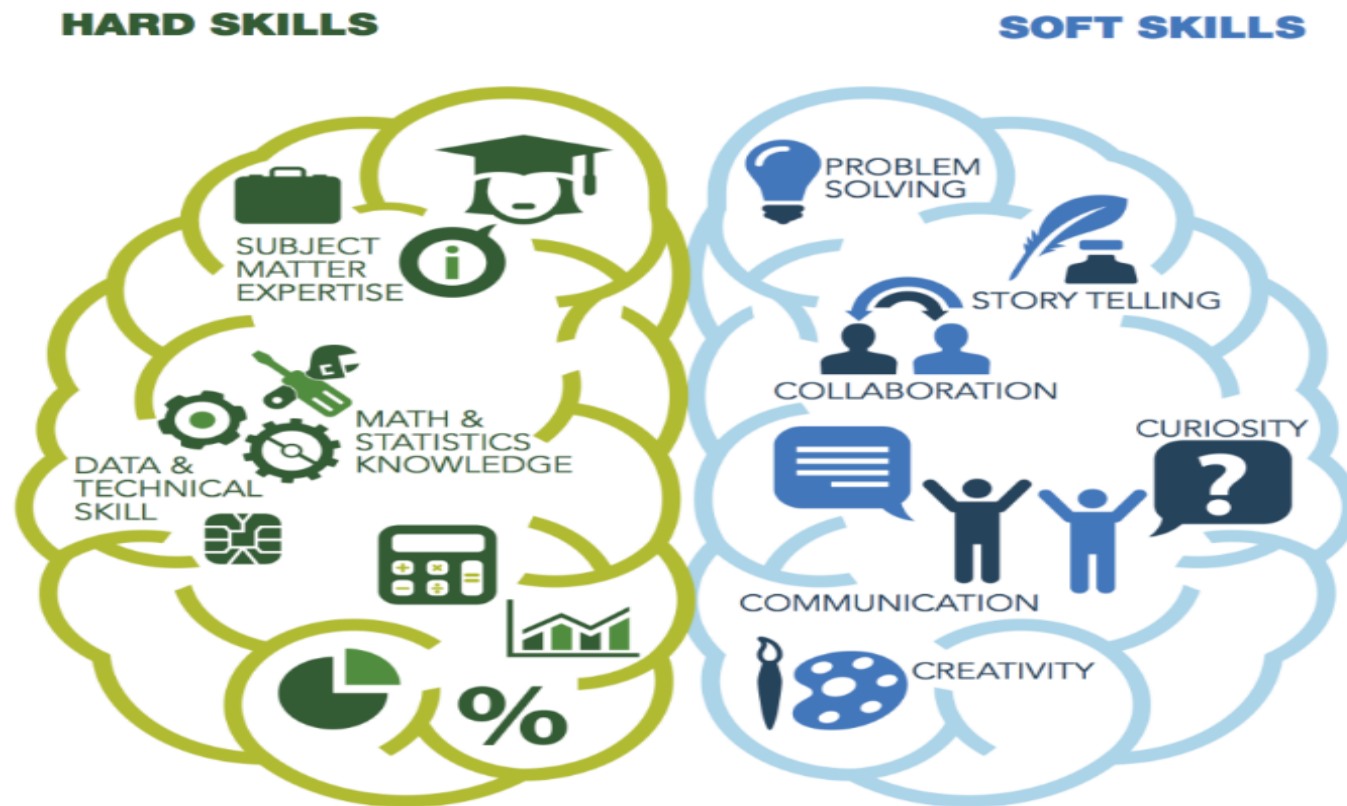
Analytics Process
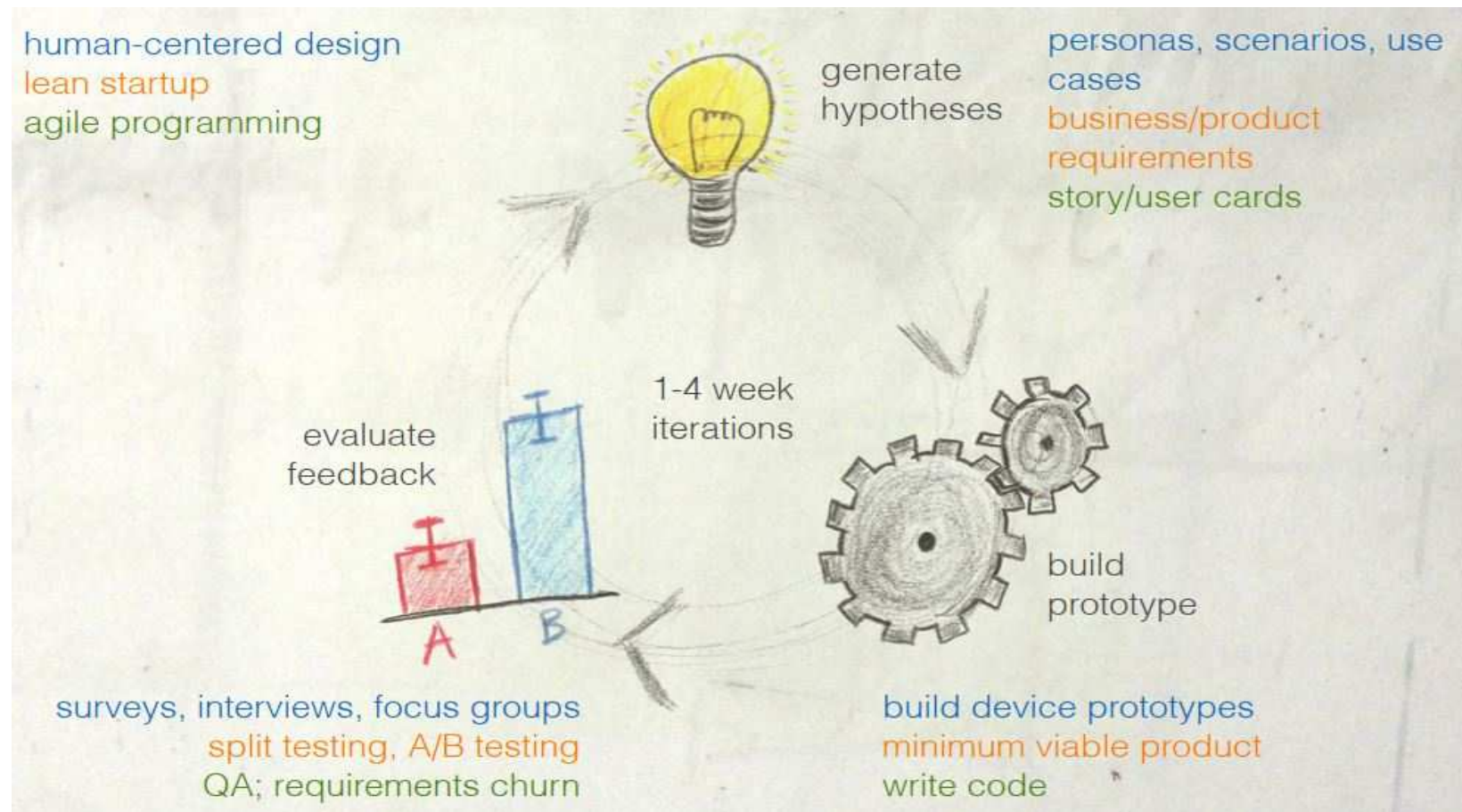
Analytics Talent

1. People who love data (slice and dice it)
2. Able to communicate effectively with people and good presentation skills.
3. Knowledge and experienced
4. A working knowledge of the most commonly used programming languages for analyzing large digital datasets
5. Insight competence

Hard vs Soft Skills

Program Training Digital Innovation & Data Analytics

Topic	Module		Duration
Introduction To Data Analytics & Innovation	Modul 1	Building Blocks	120 minutes
	Modul 2	Data Structure	120 minutes
	Modul 3	Application Structure	120 minutes
	Modul 4	Data Analysis	120 minutes
Data Visualization	Modul 5	Conceptual Framework for Data Visualization	120 minutes
	Modul 6	Basic Charting	120 minutes
	Modul 7	Intermediate Charting	120 minutes
	Modul 8	Applied Visualizations	120 minutes
Quantitative Analysis	Modul 9	Practical Statistics	120 minutes
	Modul 10	Modeling	120 minutes
	Modul 11	Data Science I	120 minutes
	Modul 12	Data Science II	120 minutes
Problem Solving	Modul 13	Problem solving Framework in Data Science	120 minutes
	Modul 14	Case 1: Driving Visual Analysis with Automobile Data	120 minutes
	Modul 15	Case 2: Stock Investment Analysis	120 minutes
	Modul 16	Case 3: Predicting Customer Churn	120 minutes
Mental Readiness	Modul 17	Pengantar Business & Data Analytics	120 minutes
	Modul 18	TBD	120 minutes
	Modul 19	TBD	120 minutes
	Modul 20	TBD	120 minutes

Teaser: Find Sample Size

$$\text{Sample size} = \dfrac{\dfrac{z^2 \times p\,(1-p)}{e^2}}{1 + \left(\dfrac{z^2 \times p\,(1-p)}{e^2 N}\right)}$$

N = population size • e = Margin of error (percentage in decimal form) • z = z-score

Population size	Confidence level = 95% Margin of error			Confidence level = 99% Margin of error		
	5%	2,5%	1%	5%	2,5%	1%
100	80	94	99	87	96	99
500	217	377	475	285	421	485
1.000	278	606	906	399	727	943
10.000	370	1.332	4.899	622	2.098	6.239
100.000	383	1.513	8.762	659	2.585	14.227
500.000	384	1.532	9.423	663	2.640	16.055
1.000.000	384	1.534	9.512	663	2.647	16.317

Terima kasih!

ASTRA HONDA MOTOR – FINANCE DEPARTMENT
JAKARTA, FEBRUARI 2019

Natanael Peranginangin
E-mail:
natanael.peranginangin@gmail.com