Using Data Science in Market Research: The Ugly Truth

OR: Things I learned in 10+ years of experience in the market research industry
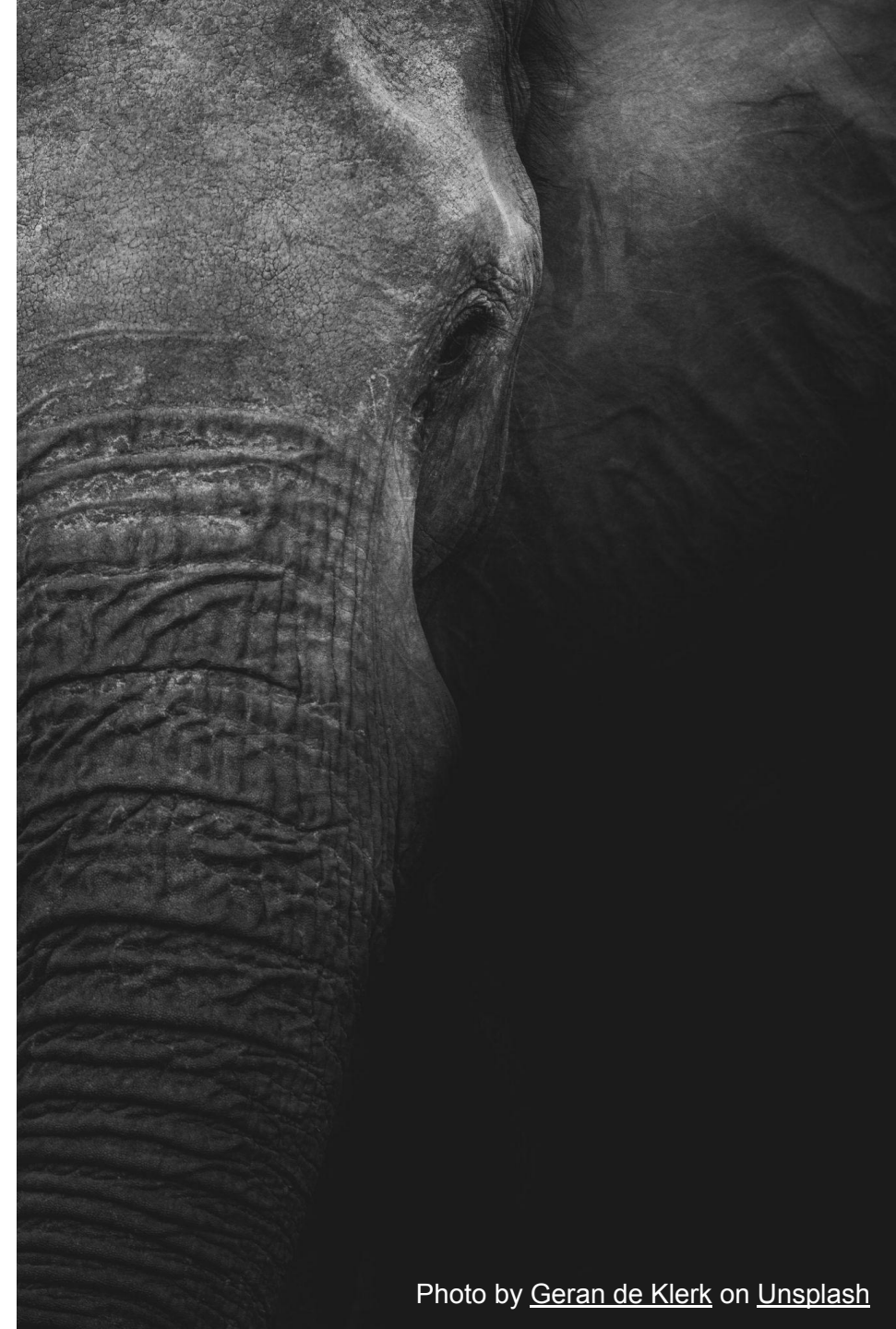
November 2022

SARID
TRUSTED INSIGHTS. PROVEN RESULTS.

The elephant in the room:
Election Surveys and Exit Polls

Predicted ~60/60 (right/left)
Outcome 64/56

Error 4/120 = 3.3%,
Which is surprisingly close to the margin of error,
but not the same thing.

$$\mathrm{MOE} = z\sqrt{\frac{p(1-p)}{500}}$$

But what they don't tell you...

SARID
TRUSTED INSIGHTS. PROVEN RESULTS.

Photo by Geran de Klerk on Unsplash

What is Market Research?
(Extract from the Wikipedia definition of market research)

- An organized effort to gather information about target markets and customers;
- A factor in maintaining competitiveness;
- Helps identify and analyze the needs of the market, the market size, and the competition;
- The systematic gathering and interpretation of information using statistical and analytical methods.

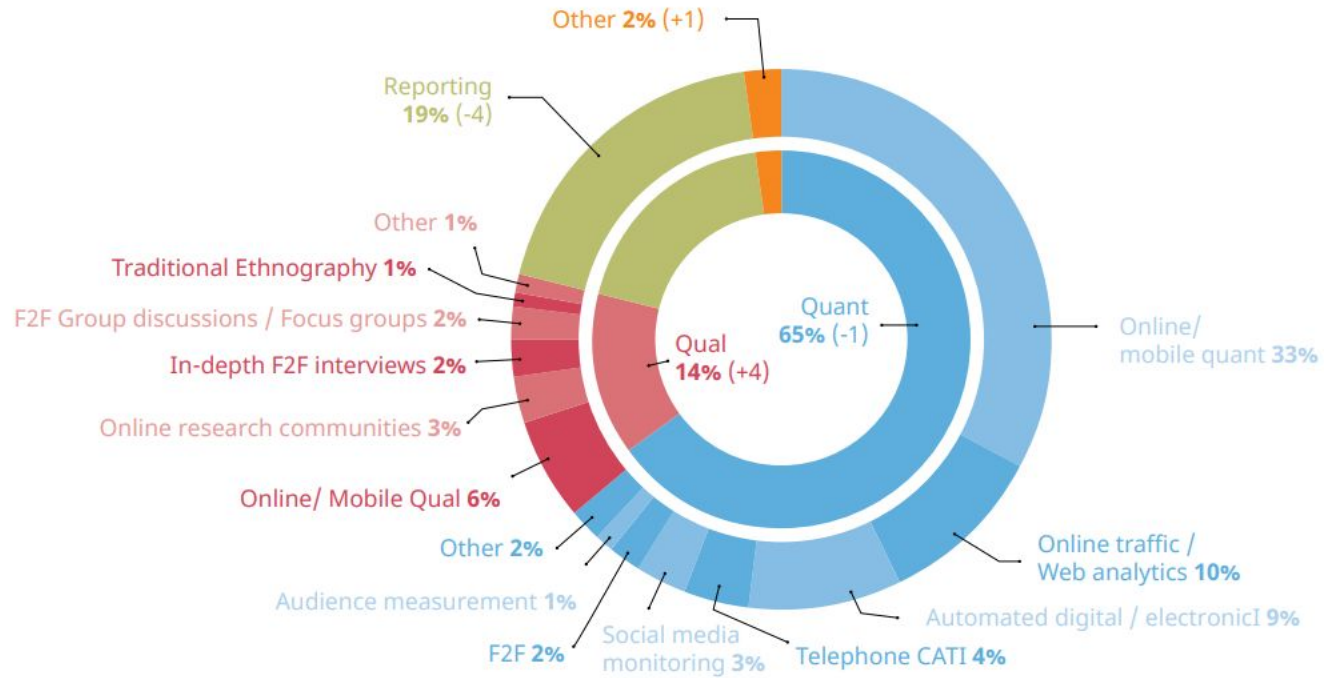- **We provide decision makers with actionable insights about their market, using scientific methods**

SARID
TRUSTED INSIGHTS. PROVEN RESULTS.

Quantitative vs. Qualitative

- **Quantitative** research - based on a "quantity of data"

  - Usually a large sample of data, i.e., hundreds of observations or more

  - Utilizing statistical tools and methodologies

  - I.e.: Surveys, data analysis of interactions/transactions, etc.

- **Qualitative** research - in depth exploration of a small sample

  - Aimed at deeper understanding of a small group of people

  - I.e.: In-depth interviews, and focus groups.

SARID
TRUSTED INSIGHTS. PROVEN RESULTS.

How market research is split between **Quantitative** and **Qualitative?**
Based on a recent ESOMAR report*

Spend by quantitative/qualitative methods

Other **2%** (+1)

Reporting **19%** (-4)

Other **1%**

Traditional Ethnography **1%**

F2F Group discussions / Focus groups **2%**

In-depth F2F interviews **2%**

Online research communities **3%**

Online/ Mobile Qual **6%**

Other **2%**

Audience measurement **1%**

F2F **2%**

Social media monitoring **3%**

Quant **65%** (-1)

Qual **14%** (+4)

Online/ mobile quant **33%**

Online traffic / Web analytics **10%**

Automated digital / electronicI **9%**

Telephone CATI **4%**

Spend per type of methodology - qualitative/quantitative
(US$ bn)

64.2

56.2

49.7

19.8

19.2

15.9

13.3

9.7

8.1

2.0

0.5

1.8

2019   2020   2021

Quantitative   Qualitative   Reporting   Other

SARID
TRUSTED INSIGHTS. PROVEN RESULTS.
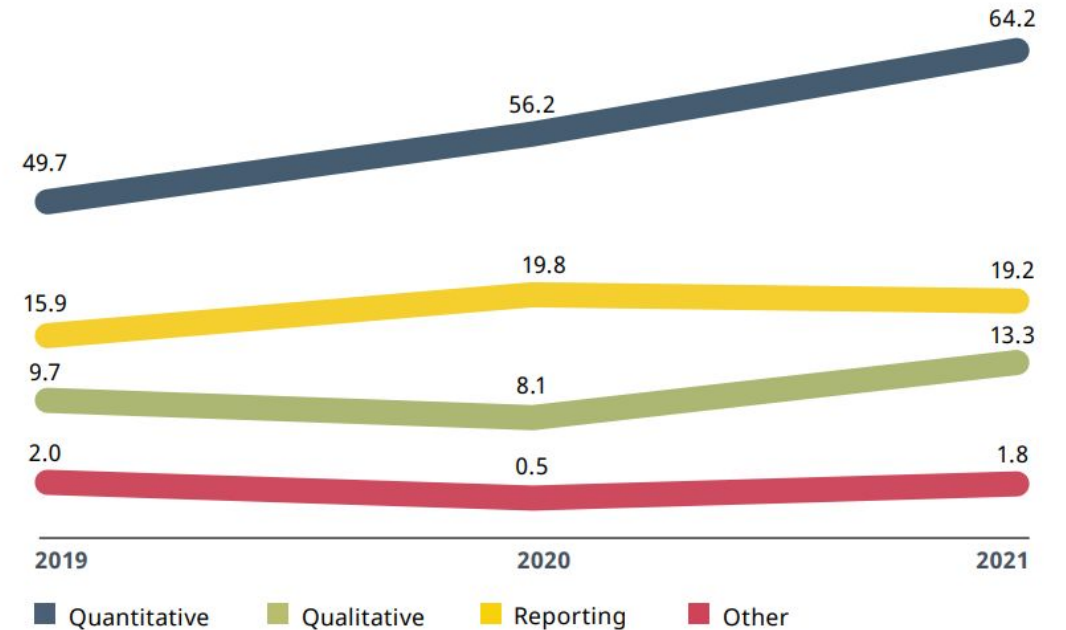
Who can benefit from Market Research?

Industry

Authorities

Hi-Tech

Retail

Services

Academia

SARID
TRUSTED INSIGHTS. PROVEN RESULTS.

How does it work?

1
Understand the bigger picture

2
Define the KPIs, the survey, and models

3
Define measurement method

4
Collect the data

5
Process the data + statistical analysis
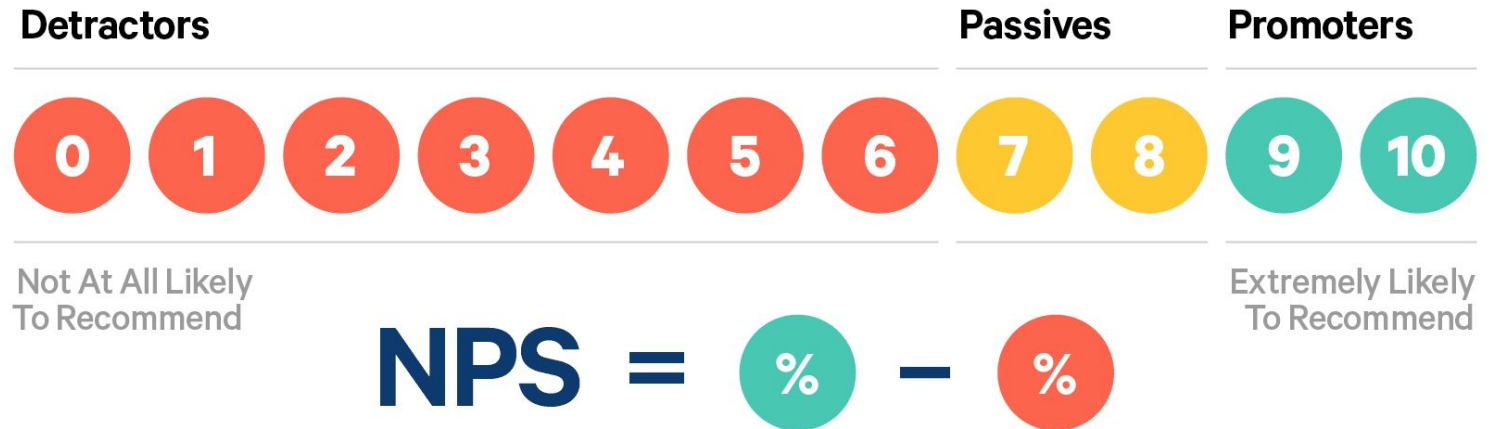
6
Results, Conclusions, Recommendations

7
Meet stakeholders, present and discuss outcomes

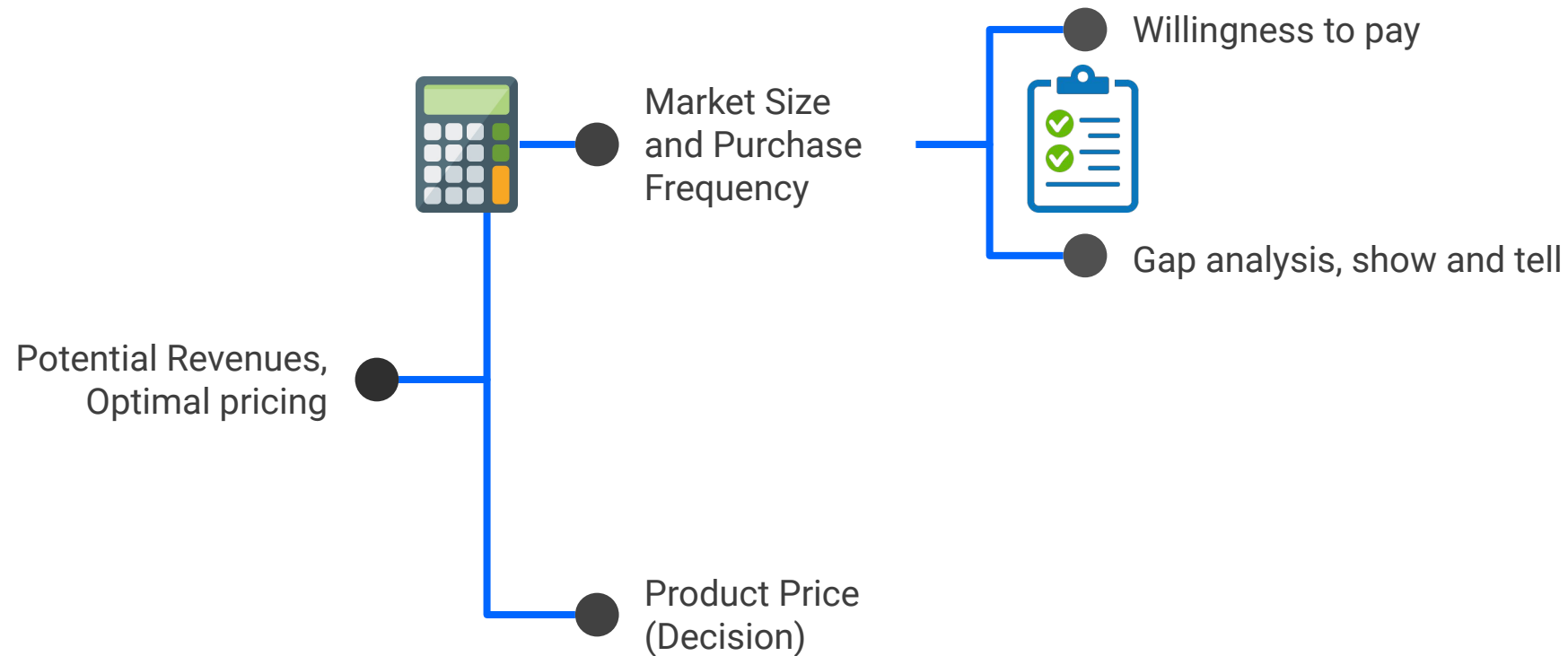SARID
TRUSTED INSIGHTS. PROVEN RESULTS.

Examples (1): VoC and NPS

VoC (=Voice of Customer) and NPS (=Net Promoter Score)

To what extent would you recommend X's services to a colleague, friend or family?   0 1 2 3 4 5 6 7 8 9 10
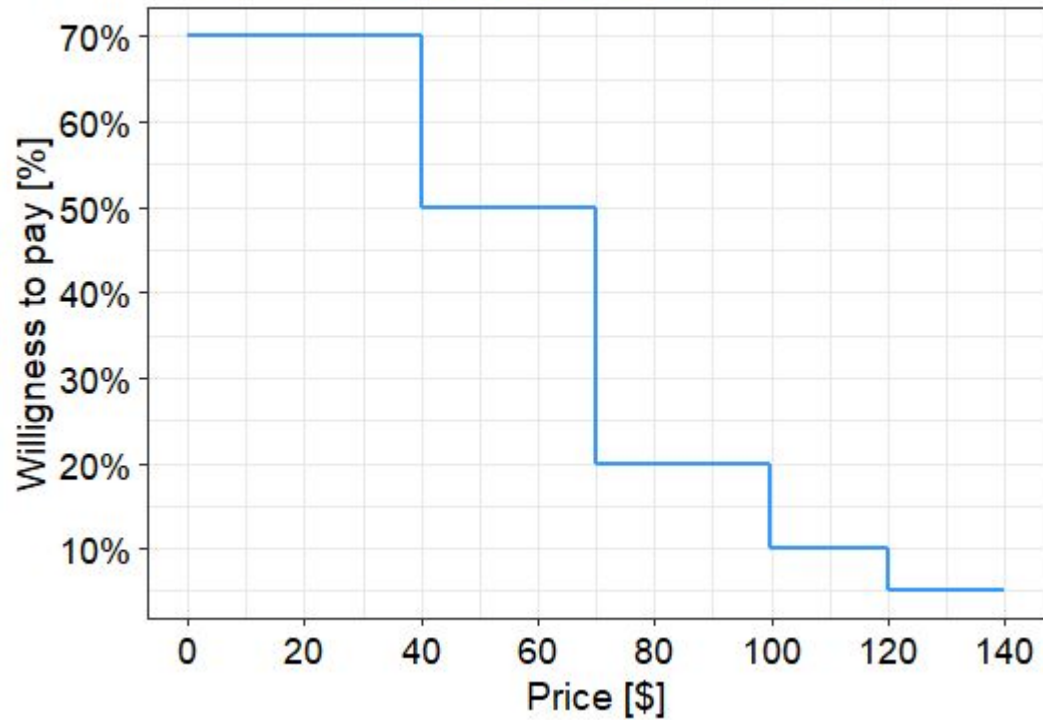
Detractors   **Passives**   **Promoters**

0 1 2 3 4 5 6   7 8   9 10

Not At All Likely
To Recommend   Extremely Likely
                                                                                To Recommend

NPS = % − %

SARID
TRUSTED INSIGHTS. PROVEN RESULTS.

Example (2): Market Size and Willingness to Pay

Willingness to pay

Market Size and Purchase Frequency

Gap analysis, show and tell

Potential Revenues, Optimal pricing

Product Price (Decision)

SARID
TRUSTED INSIGHTS. PROVEN RESULTS.

Example (2): Market Size and Willingness to Pay
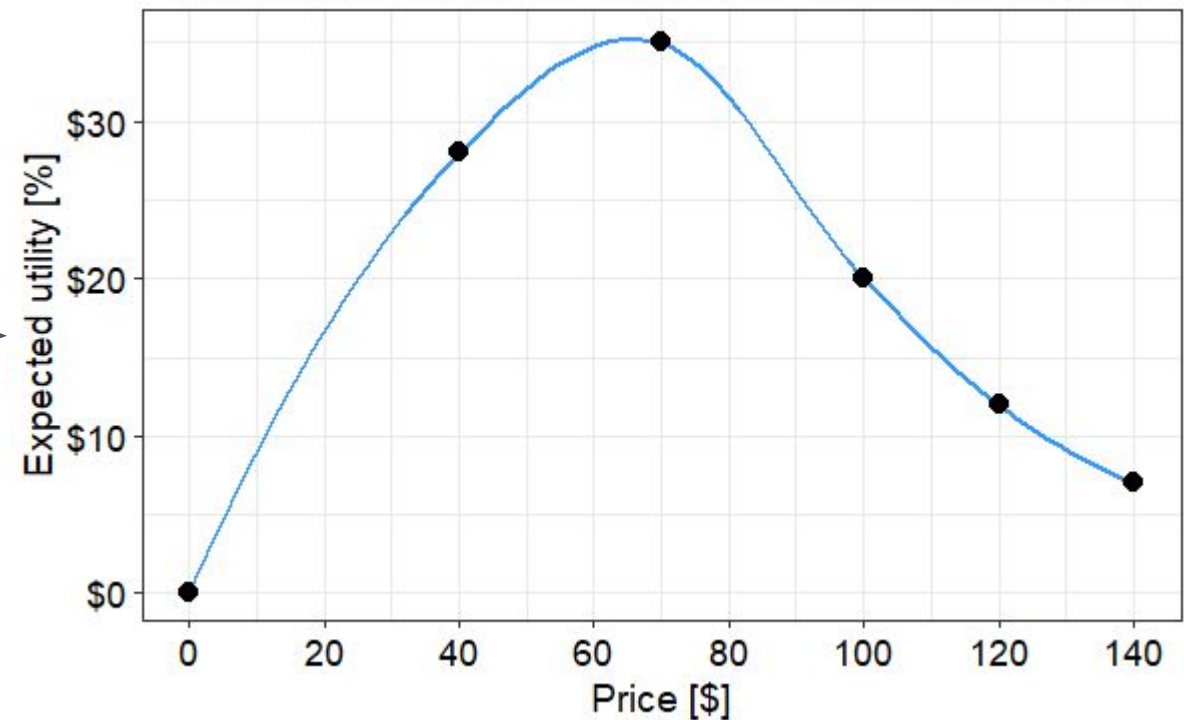
Example (3): Devising the R&D roadmap

- An existing or a new product
  - A lot of features that can be developed
  - What are the most important ones?
  - Each development has costs,
    in what should we invest?
- Use surveys to quantify the benefits
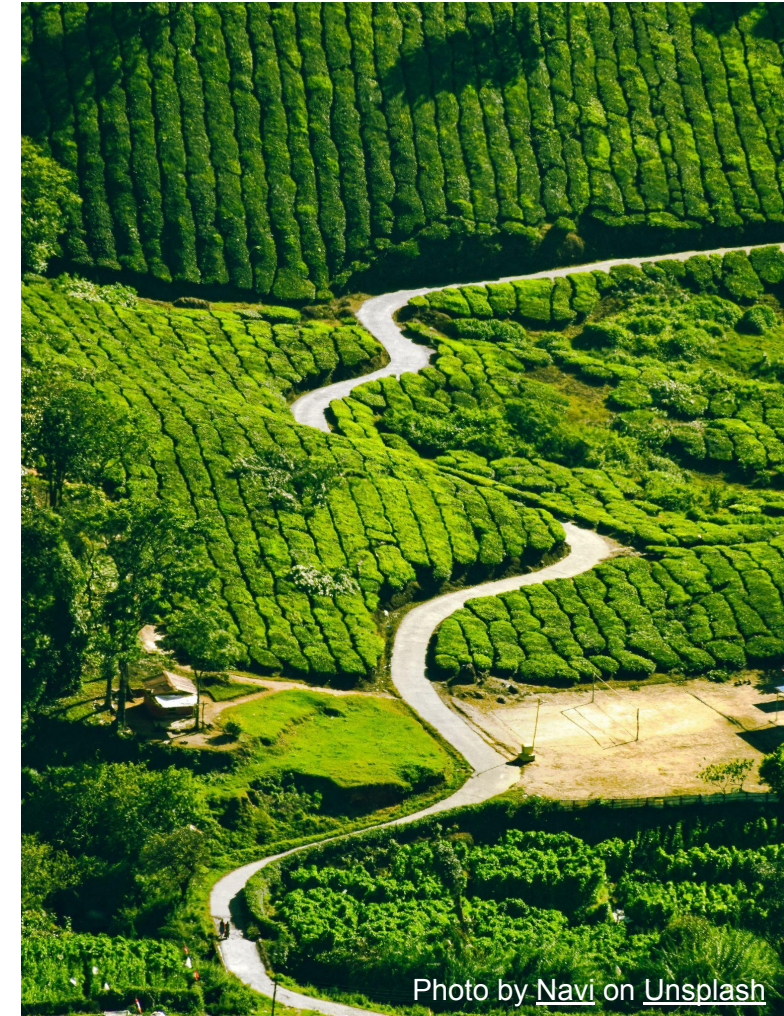- Feedback also generates feature requests
- Results help guide the R&D roadmap

Photo by Navi on Unsplash

SARID
TRUSTED INSIGHTS. PROVEN RESULTS.

Example (3): Devising the R&D roadmap

- Simulate a consumer's decision using conjoint analysis

1. In this question, we will show you four combinations of TV models. For each combination, indicate which TV model you would be most likely to purchase.

Brand	TVC	Edge	Edge	
Screen Size	40 inch	48 inch	75 inch	
Refresh Rate	240hz	120hz	120hz	None: I wouldn't choose any of these
Price	$1499	$799	$1099	
HDMI Ports	1 2.0 HDMI Port	2 2.0 HDMI Ports	3 2.0 HDMI Ports	
Wi-Fi	No Wi-Fi	Wi-Fi Enabled	No Wi-Fi	
	○	○	○	○

Next

0%

SARID
TRUSTED INSIGHTS. PROVEN RESULTS.
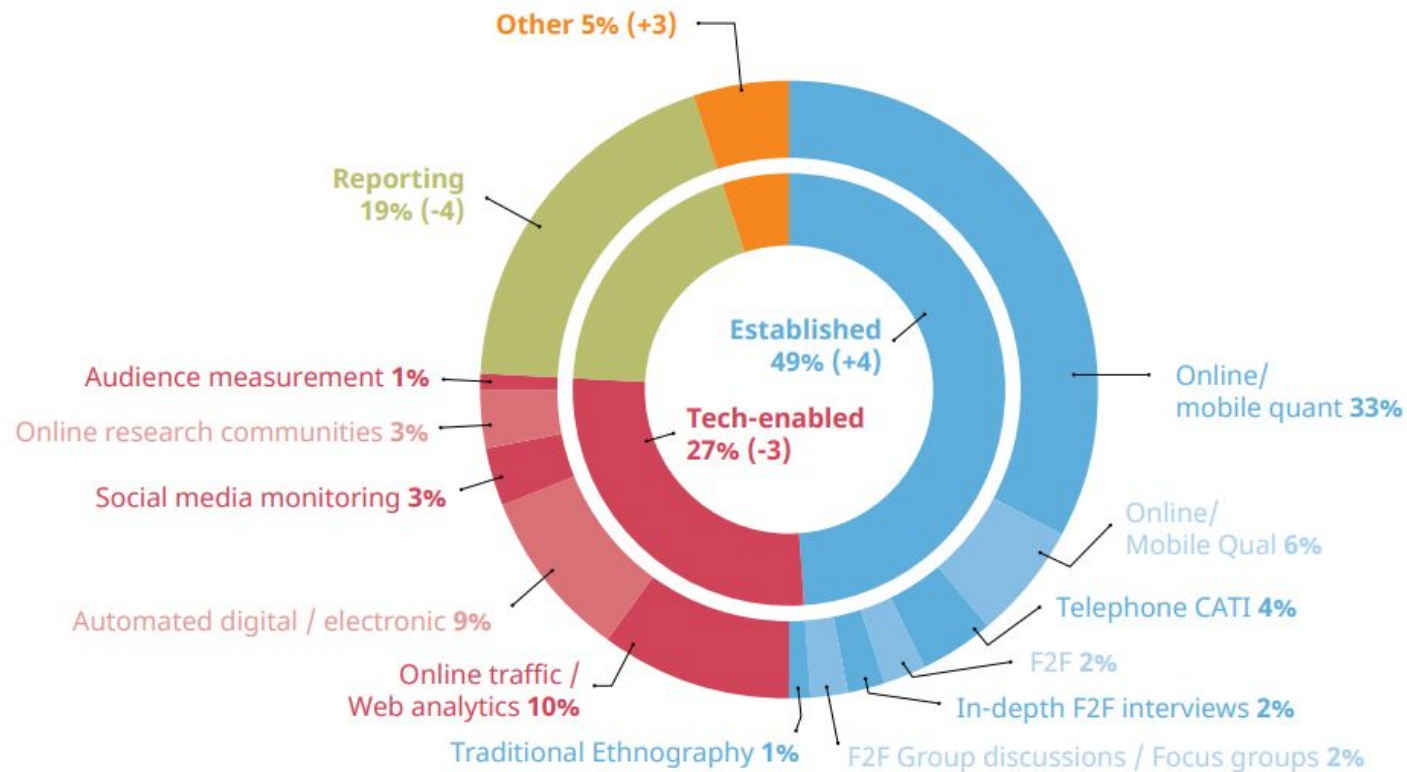
* Example adopted from the Alchemer surveying platform (link)

How market research is split between **Established** and **Tech-enabled?**
Based on a recent ESOMAR report*... (I would never use a donut chart)

Spend by established/technology-enabled methods

Other 5% (+3)

Reporting 19% (-4)

Established 49% (+4)

Tech-enabled 27% (-3)

Online/ mobile quant 33%

Audience measurement **1%**

Online research communities 3%

Social media monitoring **3%**

Automated digital / electronic **9%**

Online traffic / Web analytics **10%**

Traditional Ethnography **1%**

Online/ Mobile Qual **6%**

Telephone CATI **4%**

F2F **2%**

In-depth F2F interviews **2%**

F2F Group discussions / Focus groups **2%**

SARID
TRUSTED INSIGHTS. PROVEN RESULTS.

Survey pitfalls:
Some lessons learned the hard way

SARID
TRUSTED INSIGHTS. PROVEN RESULTS.

"This didn't teach me anything that I didn't already know"

- Some clients, rarely.

16

SARID
TRUSTED INSIGHTS. PROVEN RESULTS.

This is Nuuk
(Greenland's capital with a population of less than 20k)

SARID
TRUSTED INSIGHTS. PROVEN RESULTS.

Saving costs is more expensive...

When surveys go wrong

AKA: The "curse of dimensionality"

AKA: $n < p$

SARID
TRUSTED INSIGHTS. PROVEN RESULTS.
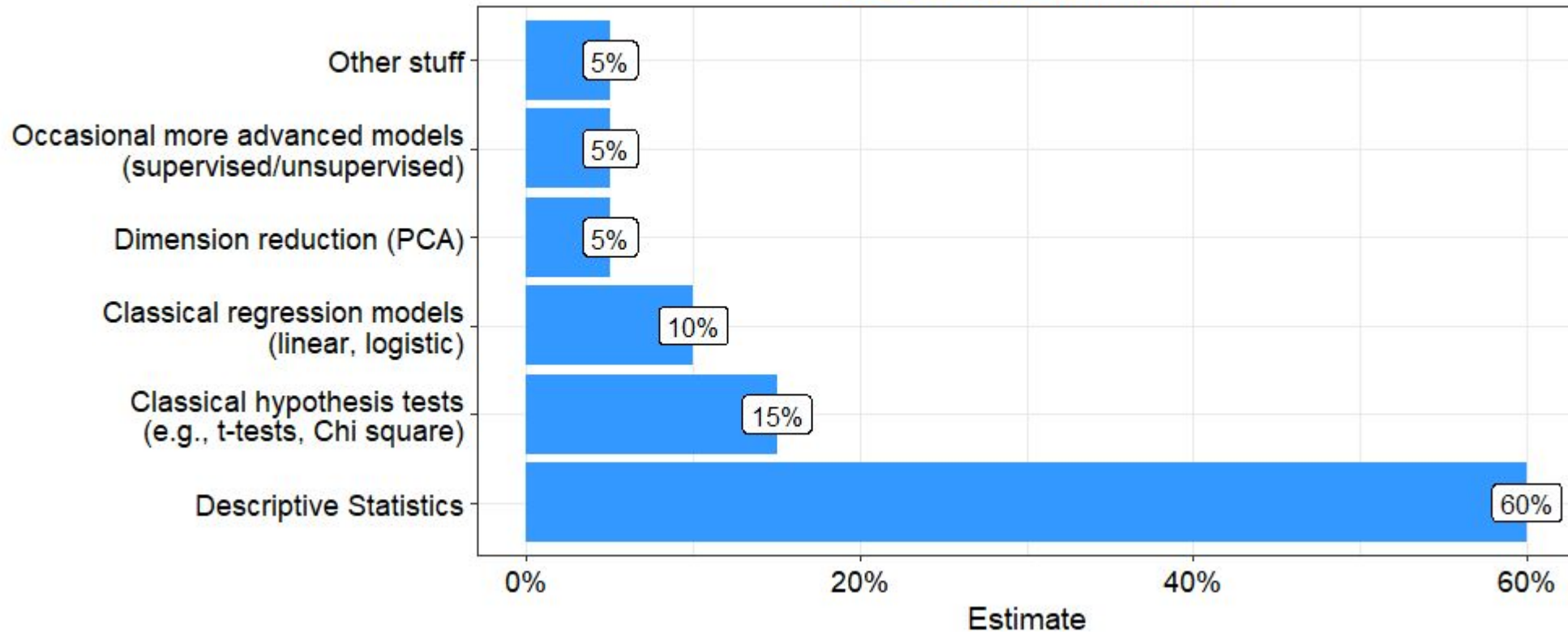
Photo by Elisa Ventur on Unsplash

Inconsistencies…
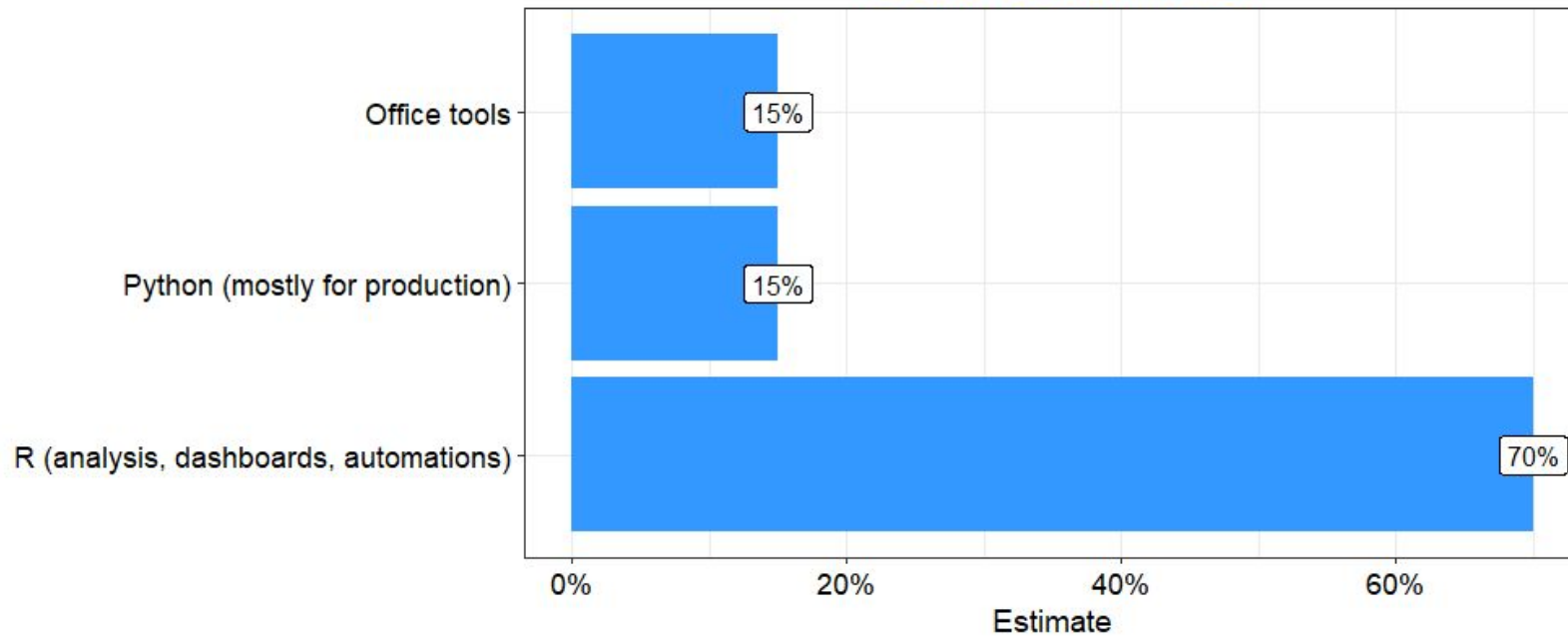
Why "the ugly truth"?

There is a lot more work to be done to incorporate data science principles into market research

What models do we employ?
(@ Sarid Research Institute)

- Other stuff — 5%
- Occasional more advanced models (supervised/unsupervised) — 5%
- Dimension reduction (PCA) — 5%
- Classical regression models (linear, logistic) — 10%
- Classical hypothesis tests (e.g., t-tests, Chi square) — 15%
- Descriptive Statistics — 60%

Estimate

SARID
TRUSTED INSIGHTS. PROVEN RESULTS.

How does our tech stack looks like?

What is our tech stack?
(@ Sarid Research Institute)

There are also additional 3rd party solutions (surveying software, git for code, etc.)

In other places in the industry you will most likely find:
SPSS / SAS, Excel, Power BI/Tableau/Qlik

To Sum Up

Market Research provides valuable information to the business that can guide strategic business decisions and day-to-day operations.

The market research industry is slowly moving towards adopting state-of-the-art methodologies and models. There is significant room for innovation.

SARID
TRUSTED INSIGHTS. PROVEN RESULTS.

Thank you for listening

Dr. Adi Sarid

@SaridResearch

https://www.sarid-ins.com

Tell me what you thought of the lecture

SARID
TRUSTED INSIGHTS. PROVEN RESULTS.