Typo Positive:
When Collection Doesn't Get You What You Need

HELLO!

I am Renee

.

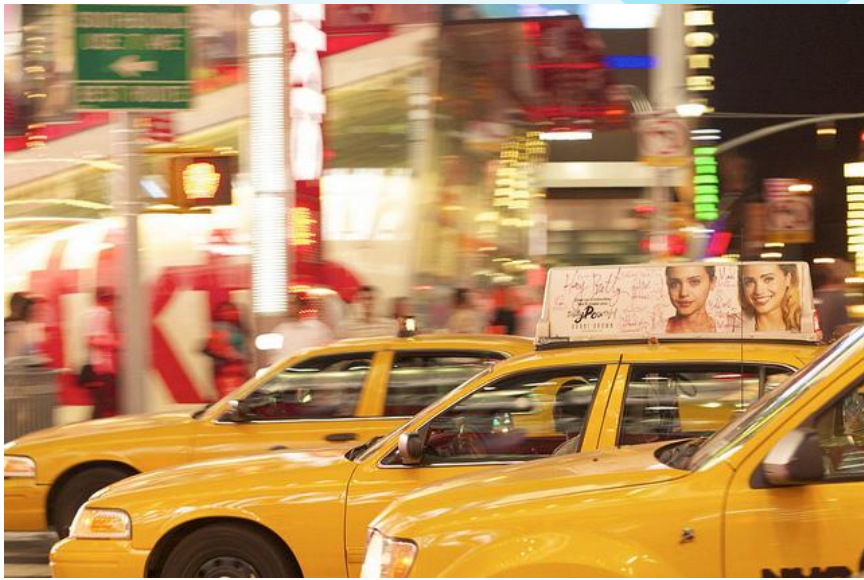
You can find me at @DataRenee

HELLO AGAIN!

I live here in NYC

.

I am a Business Intelligence Analyst



What is Data?

1

A representation of some aspect of the world

What is Data?

- Pieces of information, with a format
- Facts used for making decisions
- Information stored in and/or used by a computer

2 What is Good Data?

Fit for its intended uses in operations, planning, decision making

Why do we want good data?

- Planning
- Operations
- Look smart
- Save money
- Make decisions
- Complete transactions

3 What can we assess to check if the data might work for the intended purpose?

The Six Primary Dimensions for Data Quality Assessment

- Completeness
- Uniqueness
- Timeliness
- Validity
- Accuracy
- Consistency

PDF from UK International Data Management Association

LET'S REVIEW THESE CONCEPTS

Completeness

Are there gaps between expected data and the data we have?

Do we have enough data for the task at hand?

Uniqueness

Are there duplicates in the dataset?

Timeliness

Is there more recent data that is appropriate to the task?

Is the data accessible quickly enough?

Validity

Are the format, syntax, and type correct?

Could it be accurate?

Is the data in range?

Are there unexpected NULL values?

Accuracy

Have we stored the correct value?

Are the units of measure appropriate?

Consistency

Are there conflicts between data sources?

Does the database only change data in expected ways?

TWO KINDS OF DATA

Quantitative

Easily expressed as a number

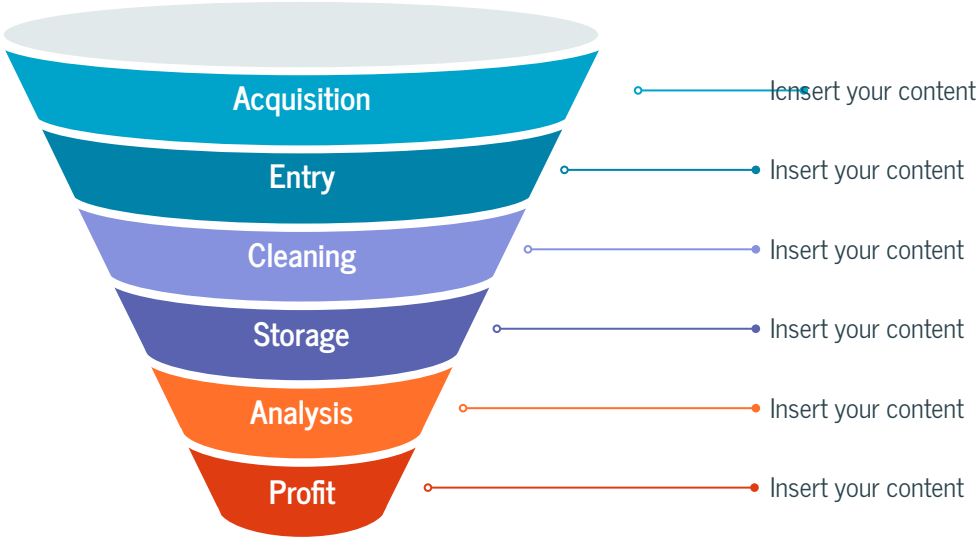
- Distance
- Weight
- Count
- Time

Qualitative

Not easily expressed as a number

- Color
- Mood
- Reason
- Opinion

DATA FUNNEL


Acquisition
Entry
Cleaning
Storage
Analysis
Profit
Insert your content
Insert your content
Insert your content
Insert your content
Insert your content
Insert your content

4 How Do We Get Data?

Collecting Data

Sensors

Questionnaires

Surveys

Interviews

Forms

Scales

Observations

Focus groups

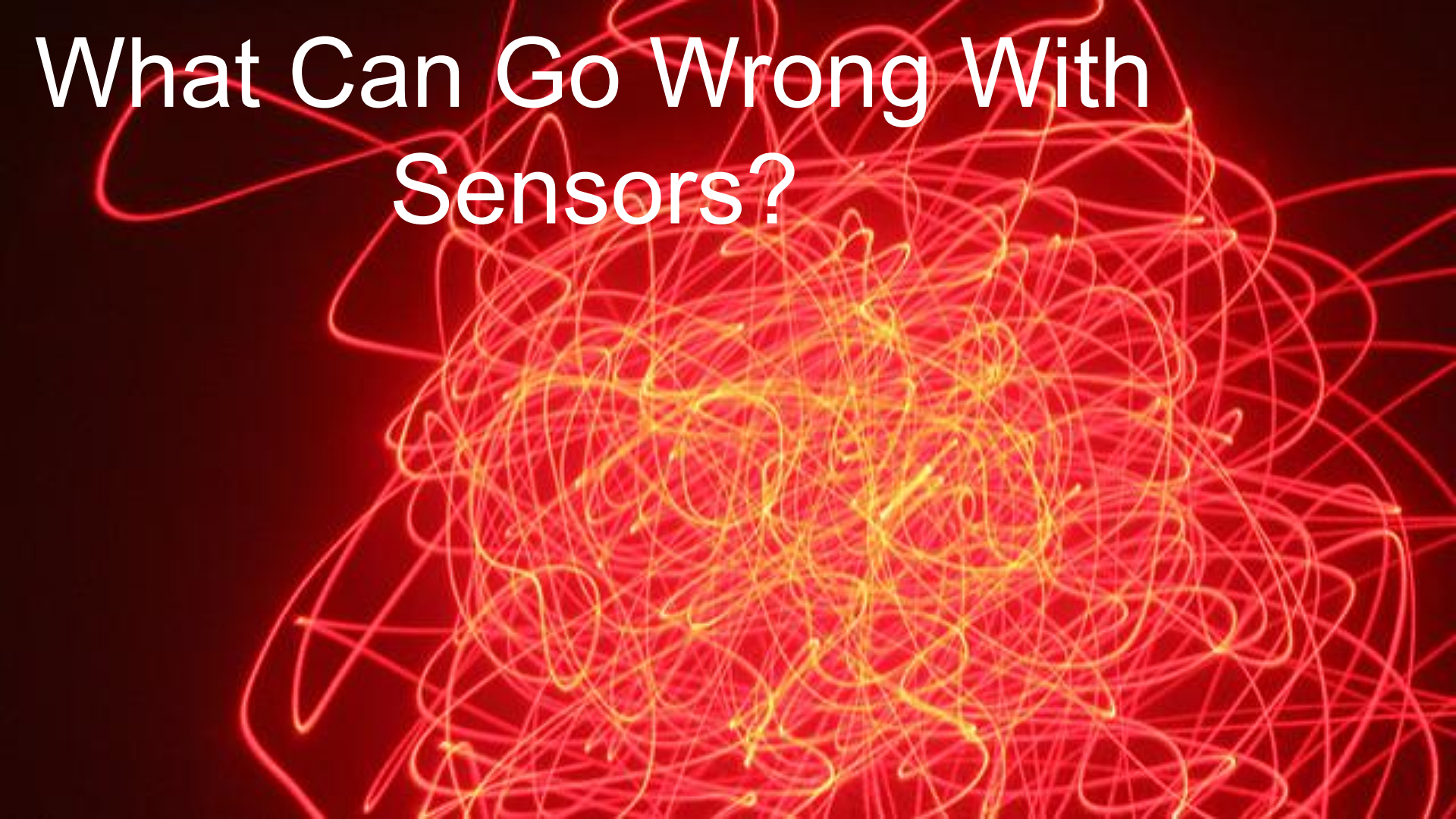
Oral histories

Archives

Clocks

Measuring sticks

What Can Go Wrong With Sensors?



SENSORS

Corrosion

Condensation (pressure drift)

Calibration

Time zone

Mounting

SENSORS

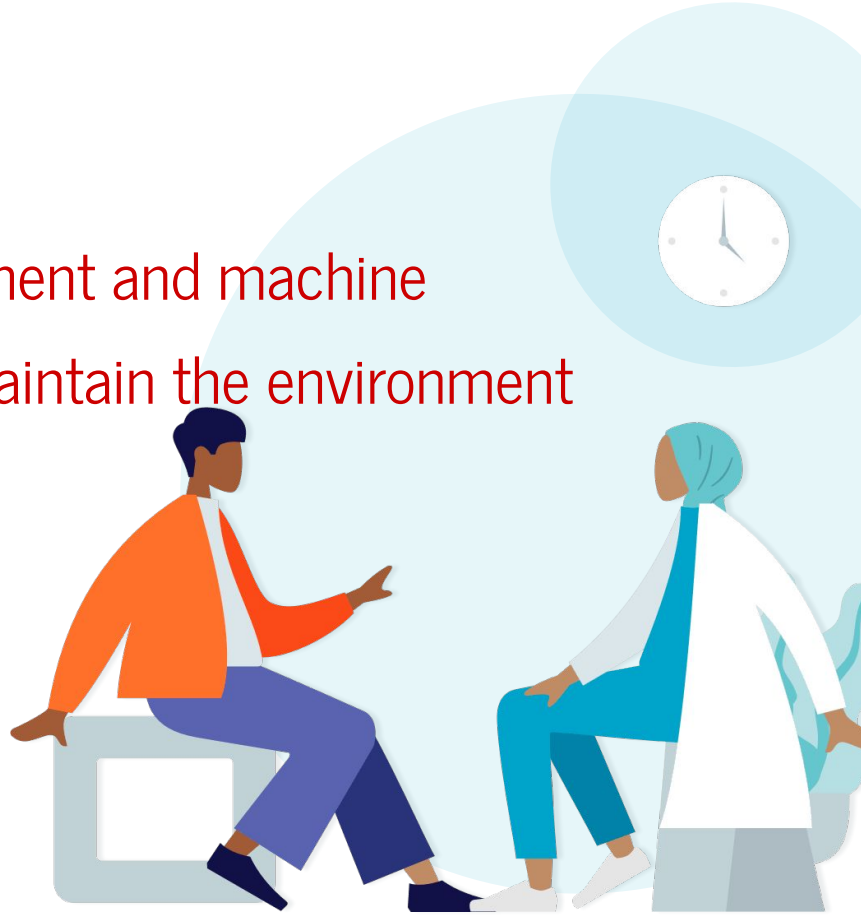
Corrosion 🔨 Maintain the environment and machine

Condensation (pressure drift) 🔨 Maintain the environment

Calibration

Time zone

Mounting



SENSORS

Corrosion

Condensation (pressure drift)

Calibration 🔨 Regular Maintenance

Time zone

Mounting 🔨 Regular Maintenance

SENSORS

Corrosion

Condensation (pressure drift)

Calibration

Time zone 🔨 Column Definition

Mounting

SENSORS

Printed Circuit Board

Cybersecurity

Fuel Quality (02)

Maintenance

SENSORS

Printed Circuit Board 🔨 Avoid 90 degree angles

Cybersecurity

Fuel Quality (02)

Maintenance

SENSORS

Printed Circuit Board

Cybersecurity 🔨 Invest in testing

Fuel Quality (02)

Maintenance

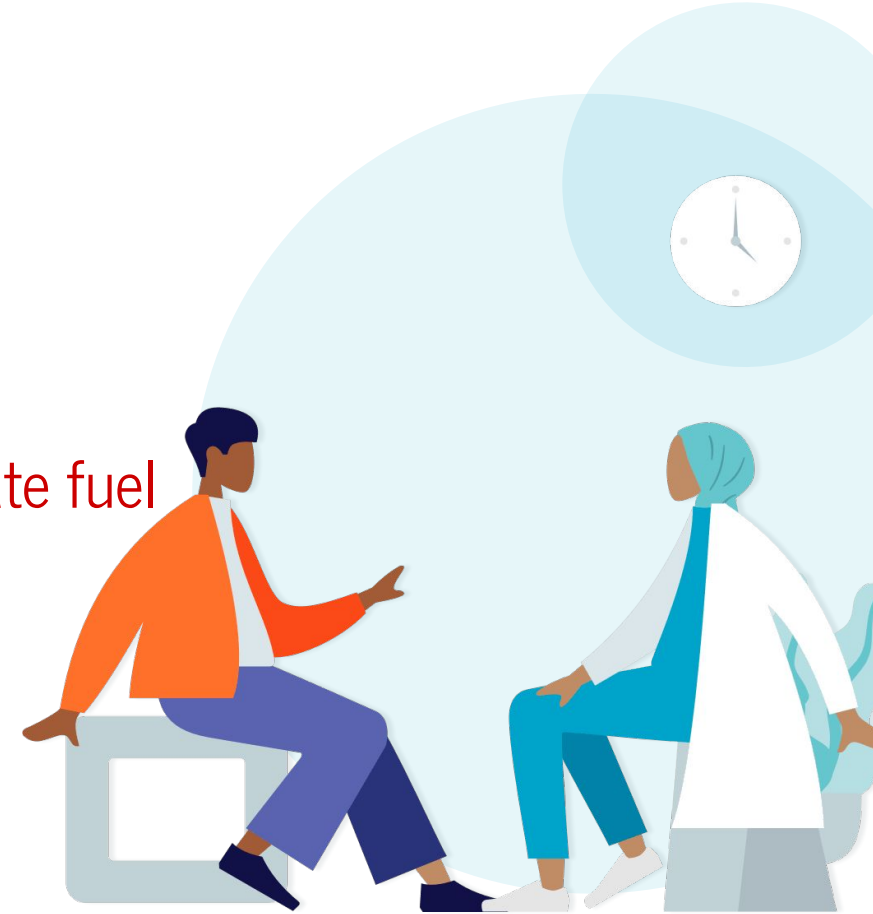
SENSORS

Printed Circuit Board

Cybersecurity

Fuel Quality (02) 🔨 Use appropriate fuel

Maintenance



SENSORS

Printed Circuit Board

Cybersecurity

Fuel Quality (02)

Maintenance 🔨 Regular Maintenance

What Can Go Wrong With GPS?

GPS

- Canyons 🔨 Sensor on top of vehicle
- Cold starts 🔨
- Not enough satellites 🔨
- Poor hardware 🔨
- Low battery 🔨

GPS

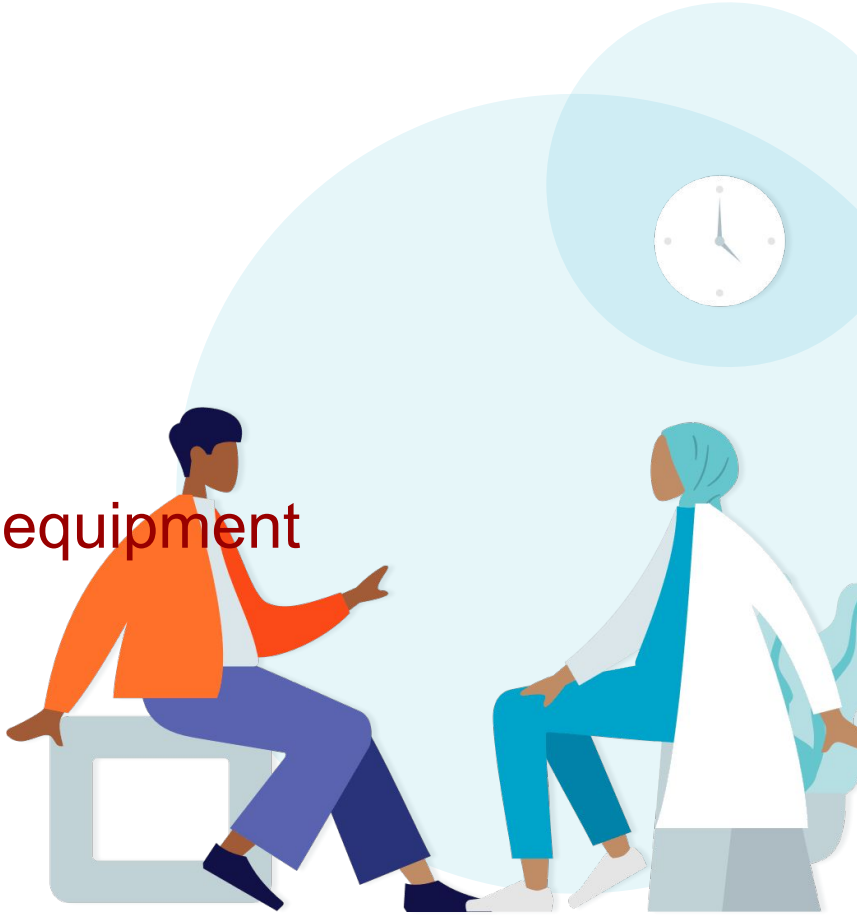
- Canyons
- Cold starts 🔨 Keep in use
- Not enough satellites
- Poor hardware
- Low battery

GPS

- Canyons
- Cold starts
- Not enough satellites 🔨 Alternate collection
- Poor hardware
- Low battery

GPS

- Canyons
- Cold starts
- Not enough satellites
- Poor hardware 🔨 Invest in equipment
- Low battery



GPS

- Canyons
- Cold starts
- Not enough satellites
- Poor hardware
- Low battery 🔨 Maintenance

GPS

Multipath signals / Bounce 🔨 Avoid urban canyons, sensor to top of vehicle

Software issues

Atmospheric conditions

GPS

Multipath signals / Bounce

Software issues 🔨 Upgrade Software

Atmospheric conditions

GPS

Multipath signals / Bounce

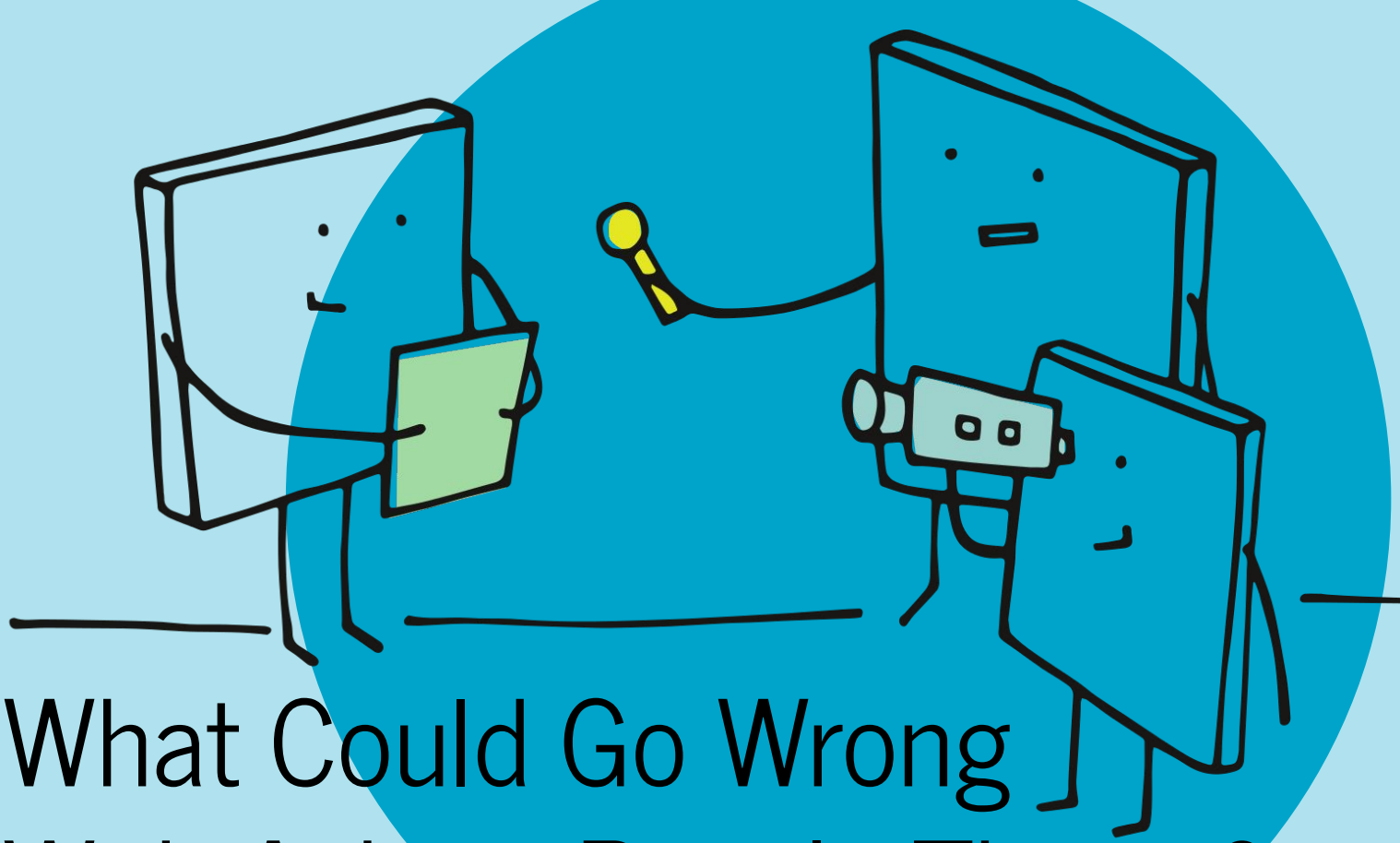
Software issues

Atmospheric conditions 🔨 Accept

Decreased collection frequency

What Can Go Wrong With Clocks?

We Literally Do Not Have Time For That



What Could Go Wrong With Asking People Things?

Questionnaires, Surveys, Focus Groups, Oral Histories, and Interviews

Choosing subjects

Cost

Contacting respondents

Keeping respondents engaged

Not flexible

Questionnaires, Surveys, Focus Groups, Oral Histories, and Interviews

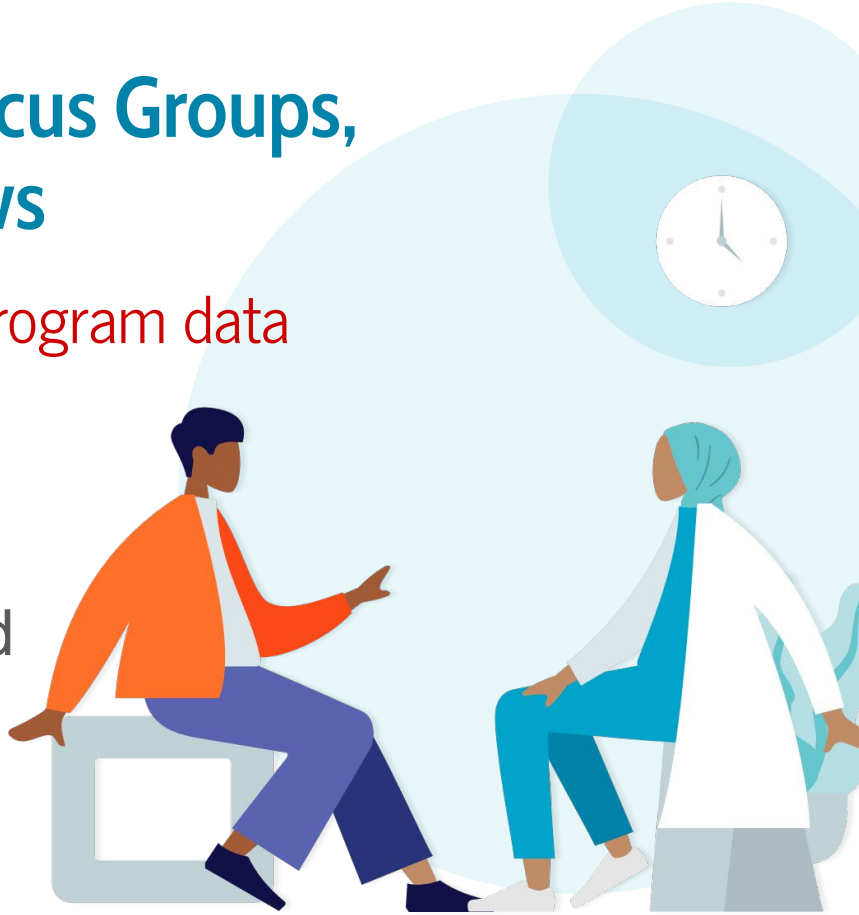
Choosing subjects 🔨 Acquire program data

Cost

Contacting Respondents

Keeping Respondents engaged

Not flexible



Questionnaires, Surveys, Focus Groups, Oral Histories, and Interviews

Choosing subjects

Cost 🔨 Assess hiring a firm vs making your own

Contacting Respondents

Keeping Respondents engaged

Not flexible

Questionnaires, Surveys, Focus Groups, Oral Histories, and Interviews

Choosing subjects

Cost

Contacting Respondents 🔨 Rely on site visitors

Keeping Respondents engaged

Not flexible

Interestingly, the first incidence of speeding flattens out at the 8 minute mark – this appears to be the level where impatience starts to set in

https://www.surveymonkey.com/curiosity/the-online-research-data-quality-problem-is-respondent-survey/

Questionnaires, Surveys, Focus Groups, Oral Histories, and Interviews

Choosing subjects

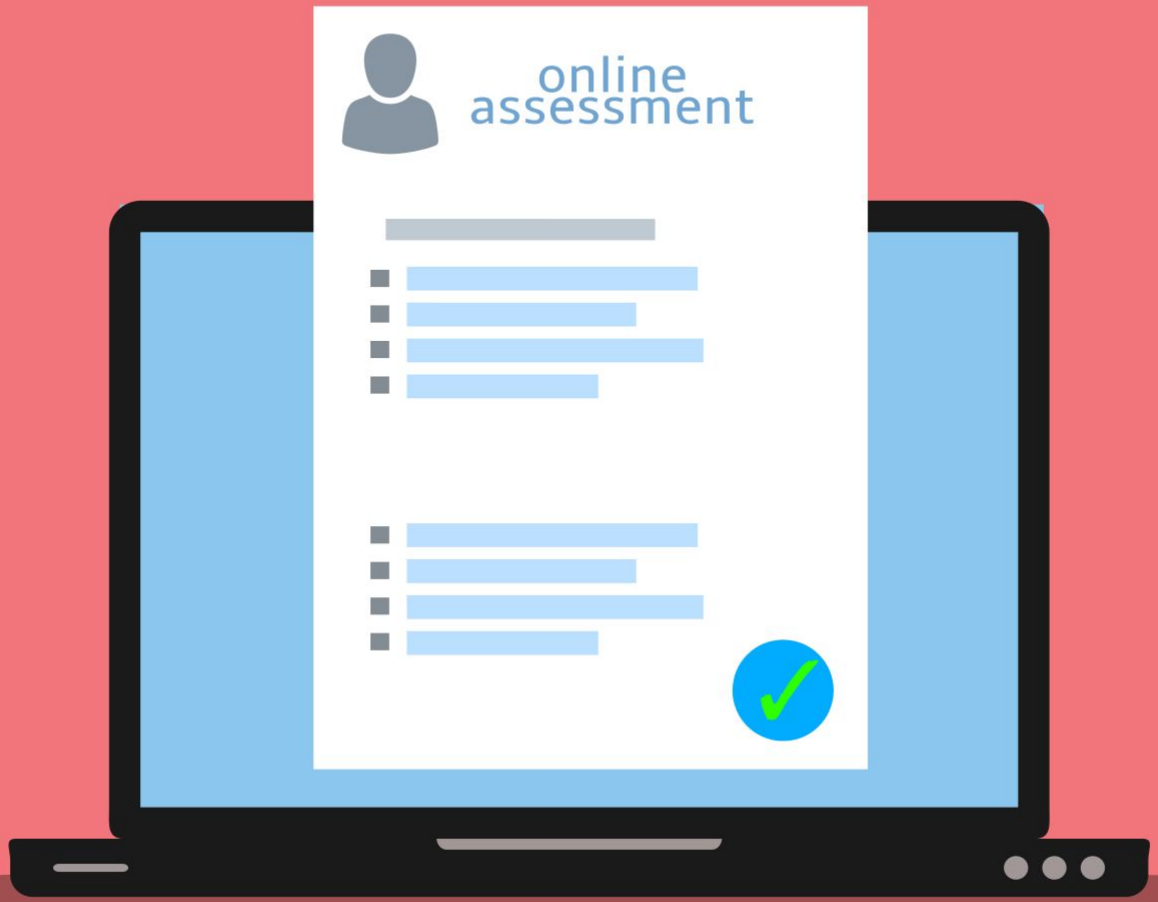
Cost

Contacting Respondents

Keeping Respondents engaged

Not flexible 🔨 Focus Groups, Oral Histories, Interviews are better

What Could Go Wrong With Forms?



Questionnaires, Surveys, Focus Groups, Oral Histories, and Interviews

Handwriting

Data entry delay

Electronic form settings

Access to technology

Questionnaires, Surveys, Focus Groups, Oral Histories, and Interviews

Handwriting 🔨 Fillable PDF (even on mobile)

Data entry delay

Electronic form settings

Access to technology

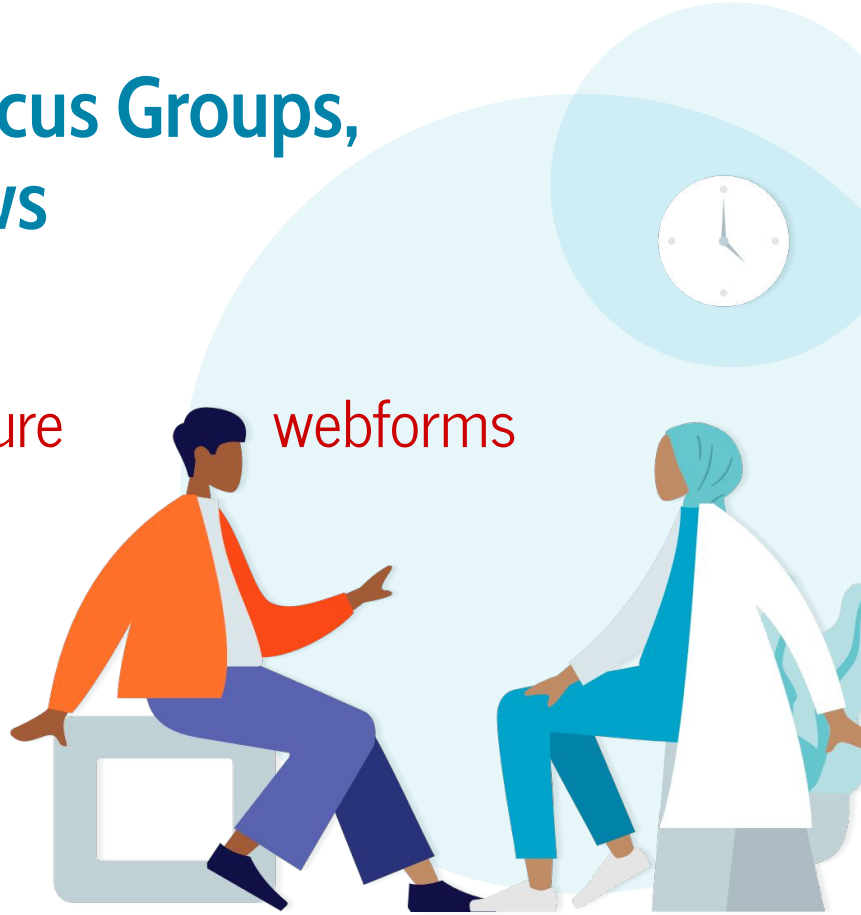
Questionnaires, Surveys, Focus Groups, Oral Histories, and Interviews

Handwriting

Data entry delay 🔨 Provide secure webforms

Electronic form settings

Access to technology



Questionnaires, Surveys, Focus Groups, Oral Histories, and Interviews

Handwriting

Data entry delay

Electronic form settings

🔨 DOUBLE CHECK THEM

Access to technology

Questionnaires, Surveys, Focus Groups, Oral Histories, and Interviews

Handwriting

Data entry delay

Electronic form settings

Access to technology

🔨 Provide technology in office

ITERATE ON COLLECTION PROCESS

Design

What data do you need?

Implement

Collect some data

Test

Revisit your needs

THANKS!

What questions do you have?

You can find me at:

- @DataRenee

CREDITS

Special thanks to all the people who made and released these awesome resources for free:

- Presentation template by SlidesCarnival
- Illustrations by Pablo Stanley
- Photographs by Unsplash