

# Google data analytics professional course

## Week-1

### Solve problems with data

#### Six problem types

Data analysts typically work with six problem types

<b>1. Making predictions</b> 	<b>2. Categorizing things</b> 	<b>3. Spotting something unusual</b> 
<b>4. Identifying themes</b> 	<b>5. Discovering connections</b> 	<b>6. Finding patterns</b> 

- *Making predictions*
- *Categorizing things*
- *Spotting something unusual*
- *Identifying themes*
- *Discovering connections*
- *Finding patterns*

## Making predictions

*A company that wants to know the best advertising method to bring in new customers is an example of a problem requiring analysts to make predictions. Analysts with data on location, type of media, and number of new customers acquired as a result of past ads can't guarantee future results, but they can help predict the best placement of advertising to reach the target audience.*

## Categorizing things

*An example of a problem requiring analysts to categorize things is a company's goal to improve customer satisfaction. Analysts might classify customer service calls based on certain keywords or scores. This could help identify top-performing customer service representatives or help correlate certain actions taken with higher customer satisfaction scores.*

## Spotting something unusual

*A company that sells smart watches that help people monitor their health would be interested in designing their software to spot something unusual. Analysts who have analyzed aggregated health data can help product developers determine the right algorithms to spot and set off alarms when certain data doesn't trend normally.*

## Identifying themes

*User experience (UX) designers might rely on analysts to analyze user interaction data. Similar to problems that require analysts to categorize things, usability improvement projects might require analysts to identify themes to help prioritize the right product features for improvement. Themes are most often used to help researchers explore certain aspects of data. In a user study, user beliefs, practices, and needs are examples of themes.*

*By now you might be wondering if there is a difference between categorizing things and identifying themes. The best way to think about it is: categorizing things involves assigning items to categories; identifying themes takes those categories a step further by grouping them into broader themes.*

## Discovering connections

*A third-party logistics company working with another company to get shipments delivered to customers on time is a problem requiring analysts to discover connections. By analyzing the wait times at shipping hubs, analysts can determine the appropriate schedule changes to increase the number of on-time deliveries.*

## Finding patterns

*Minimizing downtime caused by machine failure is an example of a problem requiring analysts to find patterns in data. For example, by analyzing maintenance data, they might discover that most failures happen if regular maintenance is delayed by more than a 15-day window.*

---

## Things to avoid when asking questions

**Leading questions:** questions that only have a particular response

- Example: *This product is too expensive, isn't it?*

*This is a leading question because it suggests an answer as part of the question. A better question might be, "What is your opinion of this product?" There are tons of answers to that question, and they could include information about usability, features, accessories, color, reliability, and popularity, on top of price. Now, if your problem is actually focused on pricing, you could ask a question like "What price (or price range) would make you consider purchasing this product?" This question would provide a lot of different measurable responses.*

**Closed-ended questions:** questions that ask for a one-word or brief response only

- Example: *Were you satisfied with the customer trial?*

*This is a closed-ended question because it doesn't encourage people to expand on their answer. It is really easy for them to give one-word responses that aren't very informative. A better question might be, "What did you learn about customer experience from the trial." This encourages people to provide more detail besides "It went well."*

**Vague questions:** questions that aren't specific or don't provide context






- Example: *Does the tool work for you?*

*This question is too vague because there is no context. Is it about comparing the new tool to the one it replaces? You just don't know. A better inquiry might be, "When it comes to data entry, is the new tool faster, slower, or about the same as the old tool? If faster, how much time is saved? If it's slower, how much time is lost?" These questions give context (data entry) and help frame responses that are measurable (time).*

## More about SMART questions

**Highly effective questions are SMART questions:**

S M A R T

				
<b>S-specific</b>	<b>M-easurable</b>	<b>A-ction-oriented</b>	<b>R-elevant</b>	<b>T-ime-bound</b>
Is the question specific? Does it address the problem? Does it have context? Will it uncover a lot of the information you need?	Will the question give you answers that you can measure?	Will the answers provide information that helps you devise some type of action plan?	Is the question about the particular problem you are trying to solve?	Are the answers relevant to the specific time being studied?

# Week-2

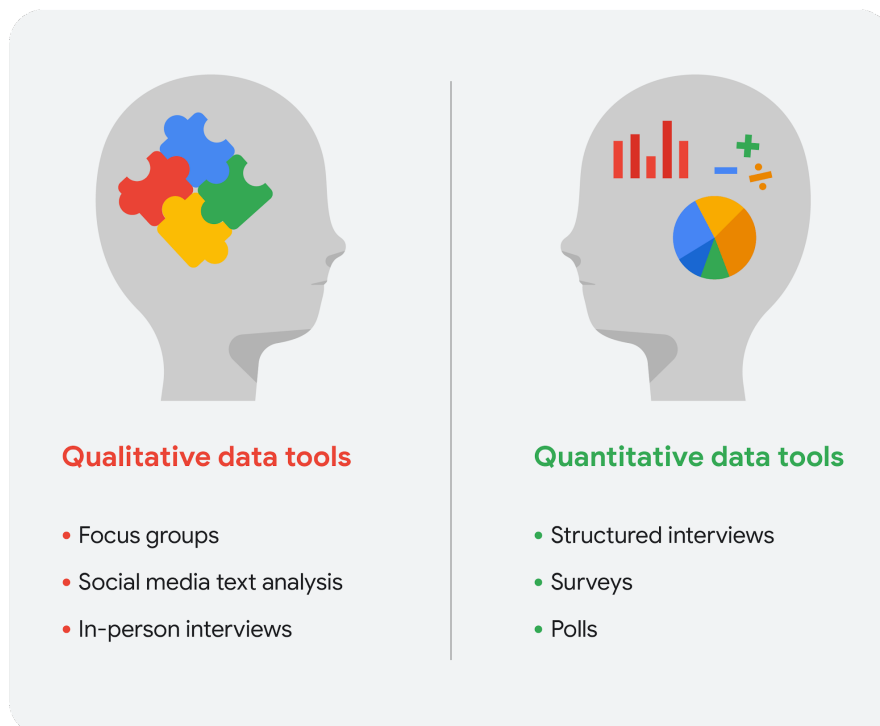
## Understand the power of data

- Data-inspired decision-making explores different data sources to find out what they have in common.
- Data derived decision making

### Quantitative and Qualitative.

**Quantitative data** is all about the specific and objective measures of numerical facts. This can often be the what, how many, and how often about a problem. In other words, things you can measure, like how many commuters take the train to work every week.

On the other hand, **qualitative data** describes subjective or explanatory measures of qualities and characteristics or things that can't be measured with numerical data, like your hair color. Qualitative data is great for helping us answer why questions.



## **Quantitative data Questions**

- *First are measurable questions.*
- *How many negative reviews are there?*
- *What's the average rating?*
- *How many of these reviews use the same keywords?*
- *These questions generate quantitative data, numerical results that help confirm their customers aren't satisfied.*

## **Qualitative data Questions**

- *Why are customers unsatisfied?*
- *How can we improve their experience?*
- *These are questions that lead to qualitative data.*

### **Strategic**

*Focuses on long term goals and strategies at the highest level of metrics.*

### **Operational**

*Short-term performance tracking and intermediate goals.*

### **Analytical**

*Consists of the datasets and the mathematics used in these sets.*

## FORMULA

*The net profit over a period of time and the cost of investment.*

*net profit over a period of time / the cost of investment*

## Follow the evidence

### Data versus metrics

**Data** is a collection of facts. **Metrics** are quantifiable data types used for measurement.

### Designing compelling dashboards

- Check the pdf name "M2\_W2\_Designing compelling dashboards".

## Connecting the data dots

### Big and small data



Small data	Big data
Describes a data set made up of specific metrics over a short, well-defined time period	Describes large, less-specific data sets that cover a long time period
Usually organized and analyzed in spreadsheets	Usually kept in a database and queried
Likely to be used by small and midsize businesses	Likely to be used by large organizations
Simple to collect, store, manage, sort, and visually represent	Takes a lot of effort to collect, store, manage, sort, and visually represent
Usually already a manageable size for analysis	Usually needs to be broken into smaller pieces in order to be organized and analyzed effectively for decision-making

## Challenges and benefits

*Here are some challenges you might face when working with big data:*

- *A lot of organizations deal with data overload and way too much unimportant or irrelevant information.*
- *Important data can be hidden deep down with all of the non-important data, which makes it harder to find and use. This can lead to slower and more inefficient decision-making time frames.*
- *The data you need isn't always easily accessible.*
- *Current technology tools and solutions still struggle to provide measurable and reportable data. This can lead to unfair algorithmic bias.*
- *There are gaps in many big data business solutions.*

*Now for the good news! Here are some benefits that come with big data:*

- *When large amounts of data can be stored and analyzed, it can help companies identify more efficient ways of doing business and save a lot of time and money.*



- *Big data helps organizations spot the trends of customer buying patterns and satisfaction levels, which can help them create new products and solutions that will make customers happy.*
- *By analyzing big data, businesses get a much better understanding of current market conditions, which can help them stay ahead of the competition.*
- *As in our earlier social media example, big data helps companies keep track of their online presence—especially feedback, both good and bad, from customers. This gives them the information they need to improve and protect their brand.*

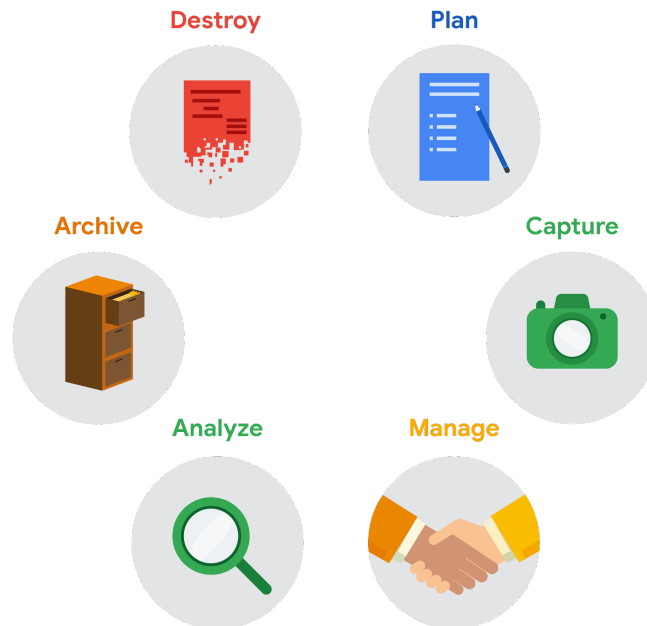
### ***The three (or four) V words for big data***

*When thinking about the benefits and challenges of big data, it helps to think about the three Vs: **volume**, **variety**, and **velocity**. Volume describes the amount of data. Variety describes the different kinds of data. Velocity describes how fast the data can be processed. Some data analysts also consider a fourth V: **veracity**. Veracity refers to the quality and reliability of the data. These are all important considerations related to processing huge, complex data sets.*

## *Week-3*

### *Working with spreadsheets*

#### *Spreadsheets and the data life cycle*



#### *Learn more about spreadsheet basics*

- **Google Sheets cheat sheet:**  
<https://support.google.com/a/users/answer/9300022/>
- **Get started with Sheets: Create and import files:**  
[https://support.google.com/a/users/answer/9300311?hl=en&ref\\_topic=9296423/](https://support.google.com/a/users/answer/9300311?hl=en&ref_topic=9296423/)
- **Sort and filter your data:**  
<https://support.google.com/docs/answer/3540681?co=GENIE.Platform%3DDesktop&hl=en>
- **Edit and format a spreadsheet:**  
<https://support.google.com/docs/answer/46973?co=GENIE.Platform%3DDesktop&hl=en&oco=0>

# Formulas in spreadsheets

- **Formula** is a set of instructions that perform a specific calculation.
- **A range** of cells is a collection of two or more cells.
- **A cell reference** is a single cell or range of cells in a worksheet that can be used in a formula.

## Quick reference: Formulas in spreadsheets

### SPREAD SHEET BASICS

- Mathematical operators      eg: + - \* /
- Auto-filling                      eg: drag
- Absolute referencing          eg: \$ symbol
- Data range                        eg: a1:c2
- Combining with functions      eg: =COUNTIF()

## More about SPREADSHEET ERRORS and fixers

Refer PDG "M2\_W3\_More about spreadsheet errors and fixes"

- #DIV/0!
- #ERROR!                      - formula mistake
- #N/A                          - value no present
- #NAME?                        - While declaring function name
- #NUM!                         - number error
- #VALUE!                      - value not present
- #REF!                         - if row deleted

# Functions in spreadsheets

*Quick reference: Functions in spreadsheets*

**Spreadsheet commands pdf**

*Refer the pdf's*

- *M2\_W3\_Spreadsheet commands 1*
- *M2\_W3\_Spreadsheet commands 2*

COUNTIF: <https://support.google.com/docs/answer/3093480?hl=en>

*Difference between formulas and functions*

- *A formula is a set of instructions used to perform a calculation using the data in a spreadsheet.*
- *A function is a preset command that automatically performs a specific process or task using the data in a spreadsheet.*

## Save time with structured thinking

**A statement of work** is a document that clearly identifies the products and services a vendor or contractor will provide to an organization. It includes objectives, guidelines, deliverables, schedule, and costs.

**A scope of work** is project-based and sets the expectations and boundaries of a project. A scope of work may be included in a statement of work to help define project outcomes.

*As a junior data analyst, It's more typical to be asked to create a scope of work than a statement of work.*

## **Scope of work (or) structural thinking**

- Deliverables
- Timeline
- Milestones
- Reports.

*Clarifying requirements and setting expectations are two of the most important parts of a project.*

## **The importance of context**

**Context** is the condition in which something exists or happens. It can turn raw data into meaningful information.



## **QUESTIONS TO ASK ABOUT DATA**

- **Who:** The person or organization that created, collected, and/or funded the data collection
- **What:** The things in the world that data could have an impact on
- **Where:** The origin of the data
- **When:** The time when the data was created or collected
- **Why:** The motivation behind the creation or collection
- **How:** The method used to create or collect it

## *Week-4*

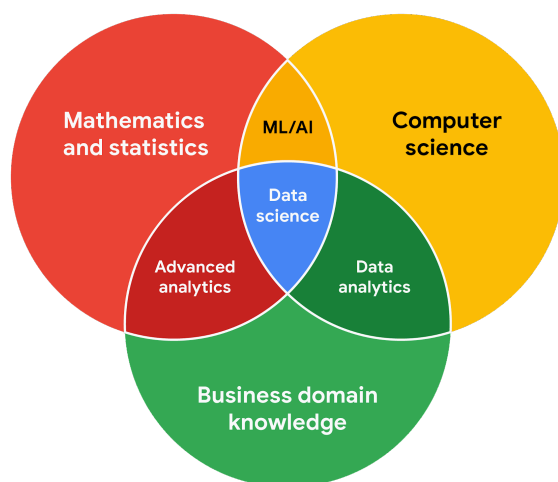
### *Balance team and stakeholder needs*

#### *Working with stakeholders*

##### **THREE TEAMS**

- *Executive team*
- *Customer-facing team*
- *Data science team*

##### *Data science team*



##### ***Working effectively with stakeholders***

- *Discuss goals*
- *Feel empowered to say "no."*
- *Plan for the unexpected*
- *Know your project*
- *Start with words and visuals*
- *Communicate often*

# Communication is key

## Data scenarios and responses

**Be sure to answer these four important questions related to your audience:**

1. *Who is your audience?*
2. *What do they already know?*
3. *What do they need to know?*
4. *How can you best communicate what they need to know?*



**Who is your audience?**



**What do they already know?**



**What do they need to know?**



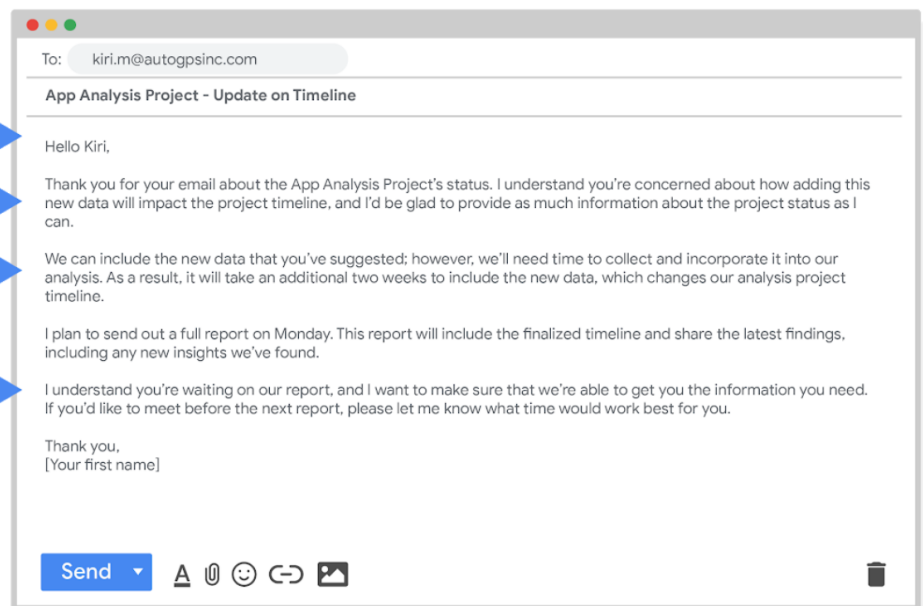
**How can you best communicate what they need to know?**

**Who** is your audience?

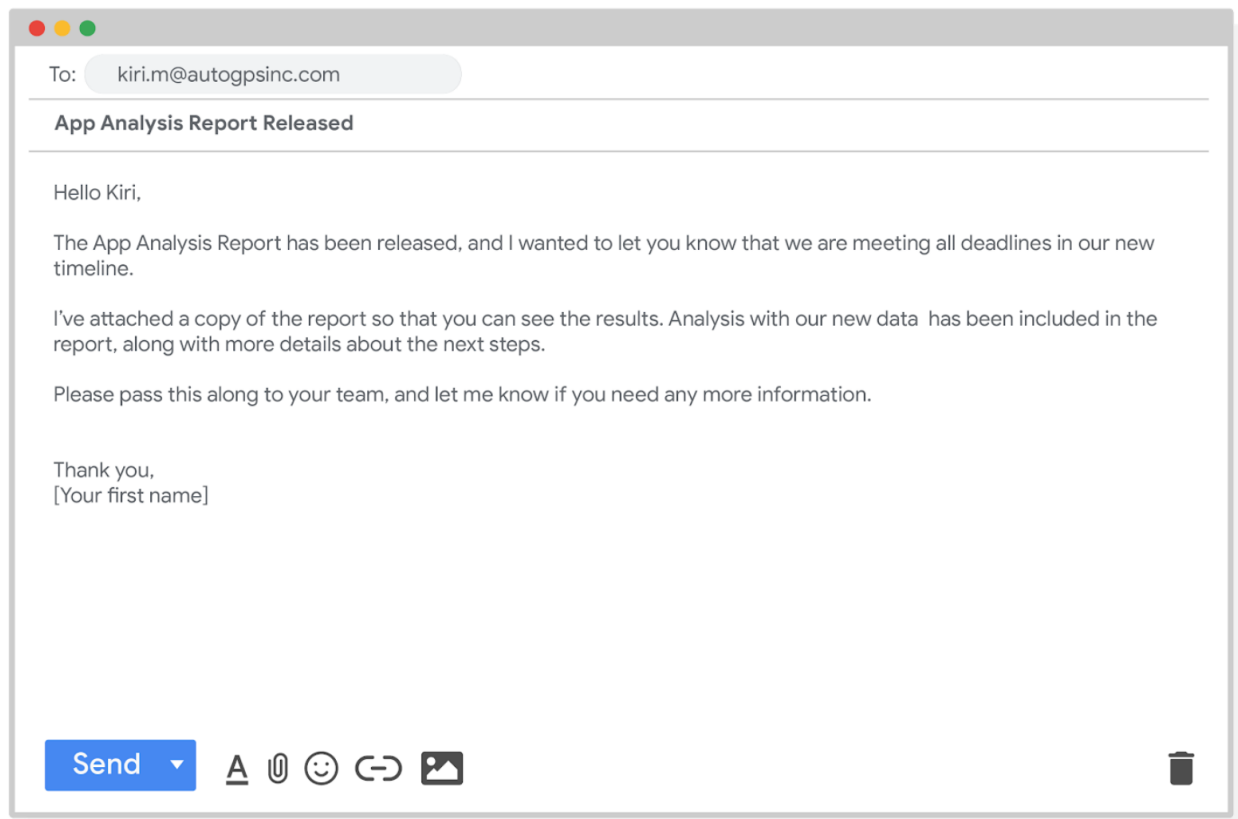
**What** do they already know?

**What** do they need to know?

**How** can you best communicate what they need to know?



*After the next report is completed, you can also send out a project update offering more information. The email could look like this:*



## *Limitations of data*



**The case of incomplete (or nonexistent!) data**



**Don't miss misaligned data**



**Deal with dirty data**



**Tell a clear story**



**Be the judge**



## **Tell a clear story**

- **Compare the same types of data:** Data can get mixed up when you chart it for visualization. Be sure to compare the same types of data and double check that any segments in your chart definitely display different metrics.
- **Visualize with care:** A 0.01% drop in a score can look huge if you zoom in close enough. To make sure your audience sees the full story clearly, it is a good idea to set your Y-axis to 0.
- **Leave out needless graphs:** If a table can show your story at a glance, stick with the table instead of a pie chart or a graph. Your busy audience will appreciate the clarity.
- **Test for statistical significance:** Sometimes two datasets will look different, but you will need a way to test whether the difference is real and important. So remember to run statistical tests to see how much confidence you can place in that difference.
- **Pay attention to sample size:** Gather lots of data. If a sample size is small, a few unusual responses can skew the results. If you find that you have too little data, be careful about using it to form judgments. Look for opportunities to collect more data, then chart those trends over longer periods.

# Amazing teamwork

## Leading great meetings

### Before the meeting

*If you are organizing the meeting, you will probably talk about the data.*

*Before the meeting:*

- *Identify your objective. Establish the purpose, goals, and desired outcomes of the meeting, including any questions or requests that need to be addressed.*
- *Acknowledge participants and keep them involved with different points of view and experiences with the data, the project, or the business.*
- *Organize the data to be presented. You might need to turn raw data into accessible formats or create data visualizations.*
- *Prepare and distribute an agenda. We will go over this next.*

### Crafting a compelling agenda

*A solid meeting agenda sets your meeting up for success. Here are the basic parts your agenda should include:*

- *Meeting start and end time*
- *Meeting location (including information to participate remotely, if that option is available)*
- *Objectives*
- *Background material or data the participants should review beforehand*

# Sample Agenda

Your name

Data Analysis Project

Phone

October 6, 2020 9:30 - 10:30 PST

Email

Group Meeting Room 1

Meeting attendees: Elon, Dae, Olivia, Kiri, Pedro

Reason for meeting: Project orientation. Set goals and draft timelines for the project.

## Goals

- Read the meeting agenda
- Review project goals
- Plan project timelines

## Questions

- Does anyone have any suggestions for the agenda?
- What sources of data have been identified and which variables will be tracked?
- What is the earliest milestone the team can schedule? What progress would the milestone mark?

## Next steps

- What should we address in the next meeting?
- 
-

## ***During the meeting***

***As the leader of the meeting, it's your job to guide the data discussion. With everyone well informed of the meeting plan and goals, you can follow these steps to avoid any distractions:***

- *Make introductions (if necessary) and review key messages*
- *Present the data*
- *Discuss observations, interpretations, and implications of the data*
- *Take notes during the meeting*
- *Determine and summarize next steps for the group*

## ***After the meeting***

***To keep the project and everyone aligned, prepare and distribute a brief recap of the meeting with next steps that were agreed upon in the meeting. You can even take it a step further by asking for feedback from the team.***

- *Distribute any notes or data*
- *Confirm next steps and timeline for additional actions*
- *Ask for feedback (this is an effective way to figure out if you missed anything in your recap)*

## ***A final word about meetings***

*Even with the most careful planning and detailed agendas, meetings can sometimes go off track. An emergency situation might steal people's attention. A recent decision might unexpectedly change requirements that were previously discussed and agreed on. Action items might not apply to the current situation. If this happens, you might be forced to shorten or cancel your meeting. That's all right; just be sure to discuss anything that impacts your project with your manager or stakeholders and reschedule your meeting after you have more information.*

# Quick Review

## Week-1

- *Six problem types*
- *Things to avoid while questioning*
- *SMART questions*

## Week-2

- *Quantitative and Qualitative data*
- *Pdf*
- *Big and small data*

## Week-3

- *Spreadsheet basics*
- *Spread formula*
- *Spreadsheet functions*
- *Structural thinking*
- *Context*

## Week-4

- *Working with stakeholders*
- *Communication*
- *Limitations of data*
- *How to perform a meeting*

DHAMODHARAN

02/10/2021