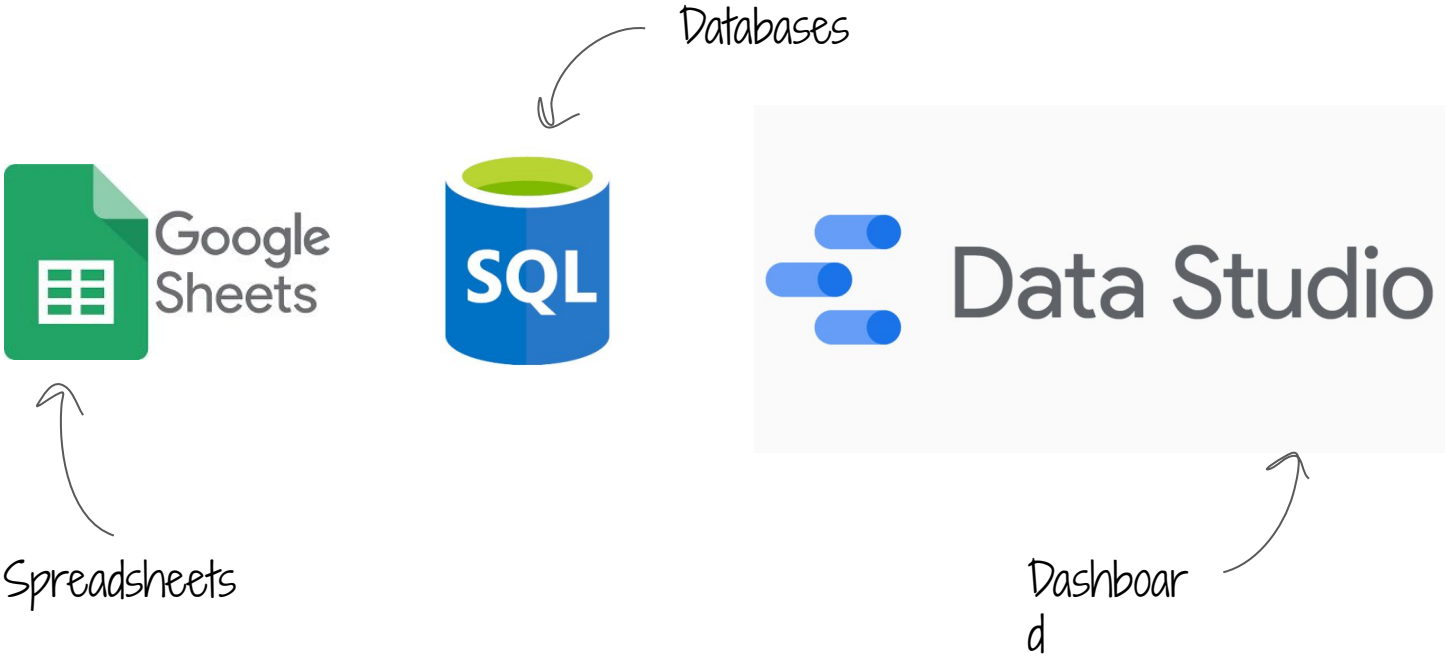
Bootcamp Live 01
Data Analytics with Google Sheets

Instructor: Kasidis Satangmongkol

Why should we learn spreadsheets?



What can spreadsheets do?

Store

Analyze

Present

Spreadsheets is basic requirement for Data Analyst

We assume you are familiar with Excel/ Sheets already

Google Sheets is free to use

fx =10+5*A4

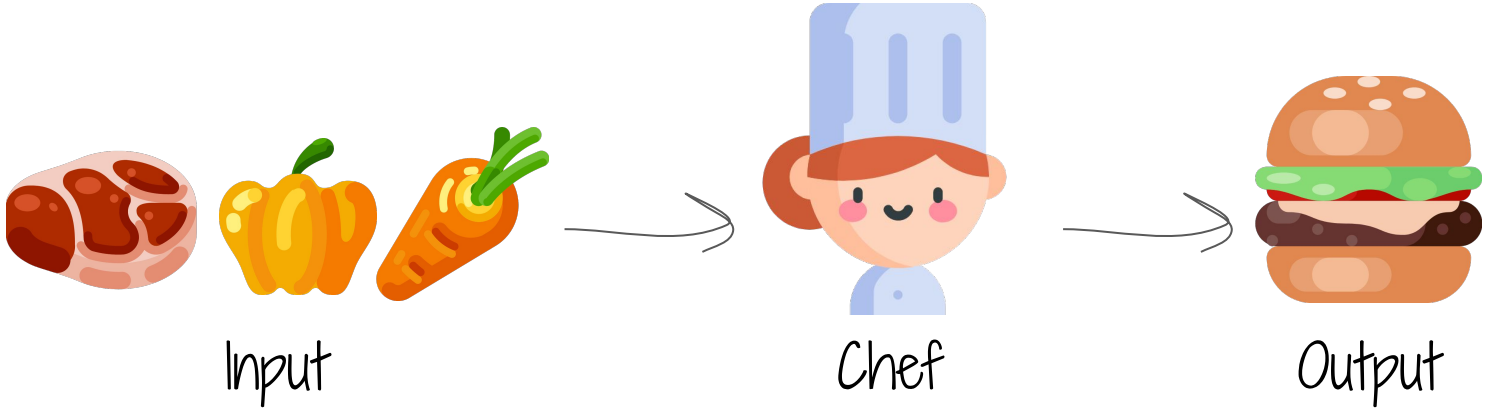
	A	B
1	**Intercept and Slope**	
2		
3	**x**	**y**
4	-10	=10+5*A4
5	-9	-35
6	-8	-30
7	-7	-25
8	-6	-20
9	-5	-15
10	-4	-10
11	-3	-5

-40 ×

Cells :)

The heart of spreadsheets is function

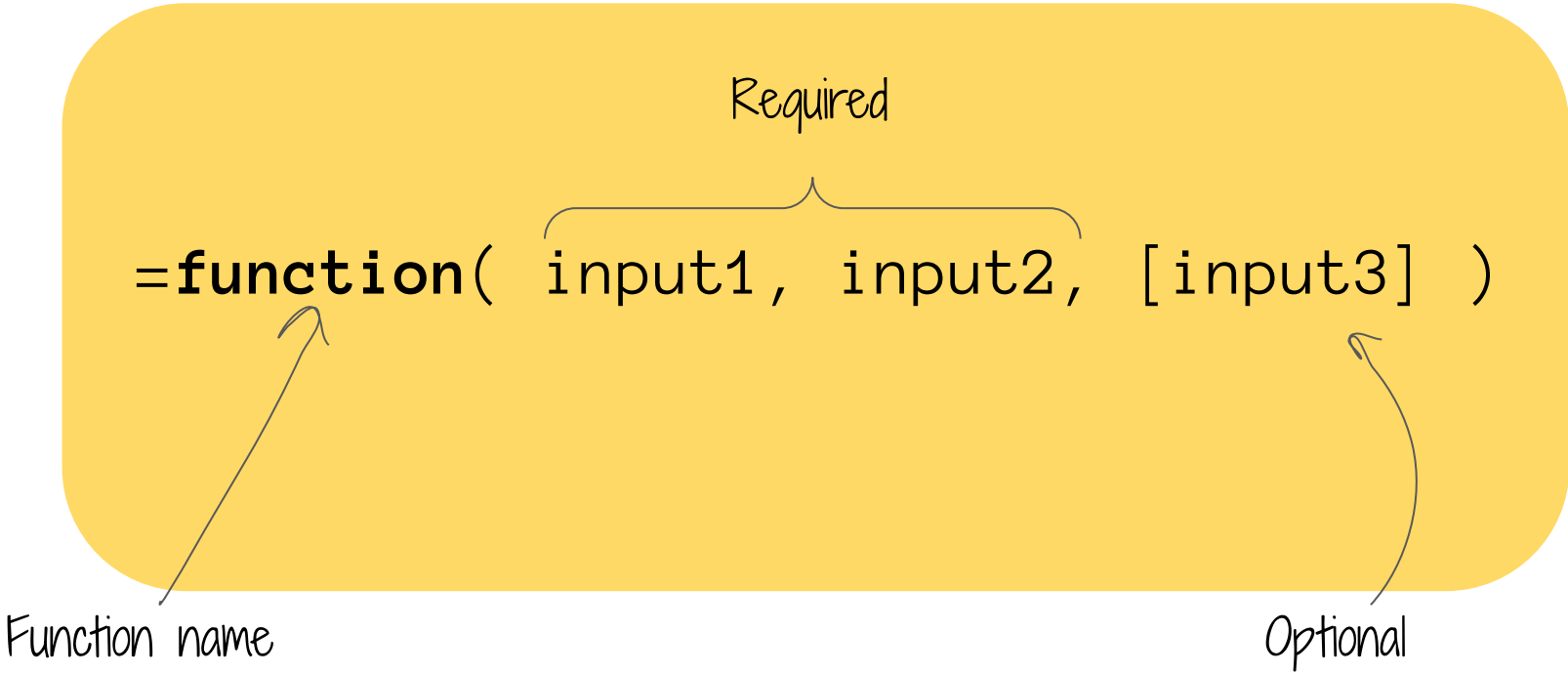
Function is your Chef



Basic Formula & Function

```
=function( input1, input2, [input3] )
```

Basic Formula & Function



The most useful function

```
=IF( score >= 80, “Passed”, “Failed”)
```

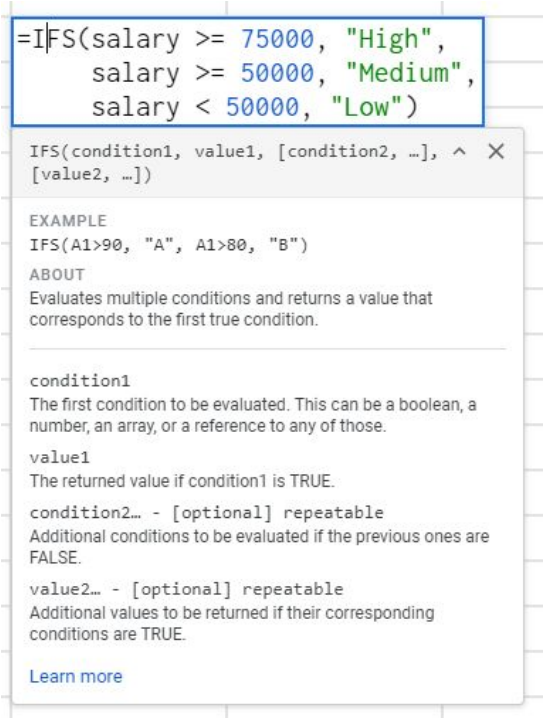
Condition TRUE FALSE

Named range

E	F	G
salary	new_salary	email
$57,500	=ArrayFormula(SALARY*1.1)	
$19,500	$21,450	
$69,000	$75,900	
$55,000	$60,500	
$47,500	$52,250	
$100,000	$110,000	
$125,000	$137,500	
$42,500	$46,750	
$69,000	$75,900	
$49,500	$54,450	
$65,000	$71,500	
$18,500	$20,350	
$65,000	$71,500	

Name Range in formula

How to write multiple lines function



ALT + ENTER

Professional Tip

Name your dataset and reference it in your formula

Example Dataset

	A	B	C	D	E	F	G
3	**ssn**	**firstname**	**lastname**	**hiredate**	**salary**	**gender**	**performance**
4	000-01-0000	Patricia	Milgrom	10/1/2004	$57,500	F	Average
5	000-02-2222	Sandy	Adams	1/15/2001	$19,500	F	Average
6	109-87-6543	Emily	Wood	3/12/1997	$69,000	F	Average
7	109-87-6544	Harold	Foster	8/14/2005	$55,000	M	Good
8	111-12-1111	James	Johnson	5/3/1996	$47,500	M	Good
9	123-45-6789	Tracy	Coulter	2/14/1993	$100,000		Good
10	222-23-2222	Bill	Marlin	3/28/1977	$125,000	M	
11	222-52-5555	Mary	Smith	1/1/2006	$42,500	F	Average
12	245-67-8910	Sandy	Johanson	6/2/2005	$69,000	F	
13	333-34-3333	Emily	Manin	12/1/2000	$49,500	F	Average
14	333-43-4444	Frank	Smith	1/29/1991	$65,000	M	Good
15	333-66-1234	Marietta	Brown	3/7/2001	$18,500	F	Poor
16	335-55-5533	Holly	Jones	4/8/1986	$65,000	F	Good
17	432-19-8765	Paul	Bronson	11/20/2003	$58,000	M	Good
18	444-45-4444	Vernon	Frank	4/10/1985	$75,000	M	Good
19	464-64-4466	David	Webster	1/29/1991	$58,500	M	Poor
20	500-50-0505	Jose	Rodriguez	7/16/1998	$150,000	M	Good
21	555-22-3333	Patricia	Rubin	7/25/2003	$45,000	F	Average
22	555-56-5555	Kenneth	Charles	6/18/1998	$40,000	M	Poor
23	612-99-1111	Melissa	Roberts	5/14/1984	$79,000	F	Good
24	625-62-6262	Holly	Holmes	6/15/1992	$55,000	F	Average
25	767-74-7373	William	Martin	8/26/2006	$23,000	M	Good
26	776-67-6666	David	Adamson	10/4/2002	$52,000	M	Poor
27	777-78-7777	Kelly	Marder	9/25/1997	$38,500	F	Average
28	925-45-7116	David	Whitehead	7/25/1980	$175,000	M	Good

Ask these questions

1. How many columns?
2. How many rows?
3. Is our data complete?

Filter data

	A	B	C	D	E	F	G
3	**ssn**	**firstname**	**lastname**	**hiredate**	**salary**	**gender**	**performance**
4	000-01-0000	Patricia	Milgrom	10/1/2004	$57,500	F	Average
5	000-02-2222	Sandy	Adams	1/15/2001	$19,500	F	Average
6	109-87-6543	Emily	Wood	3/12/1997	$69,000	F	Average
7	109-87-6544	Harold	Foster	8/14/2005	$55,000	M	Good
8	111-12-1111	James	Johnson	5/3/1996	$47,500	M	Good
9	123-45-6789	Tracy	Coulter	2/14/1993	$100,000		Good
10	222-23-2222	Bill	Marlin	3/28/1977	$125,000	M	
11	222-52-5555	Mary	Smith	1/1/2006	$42,500	F	Average
12	245-67-8910	Sandy	Johanson	6/2/2005	$69,000	F	
13	333-34-3333	Emily	Manin	12/1/2000	$49,500	F	Average
14	333-43-4444	Frank	Smith	1/29/1991	$65,000	M	Good
15	333-66-1234	Marietta	Brown	3/7/2001	$18,500	F	Poor
16	335-55-5533	Holly	Jones	4/8/1986	$65,000	F	Good
17	432-19-8765	Paul	Bronson	11/20/2003	$58,000	M	Good
18	444-45-4444	Vernon	Frank	4/10/1985	$75,000	M	Good
19	464-64-4466	David	Webster	1/29/1991	$58,500	M	Poor
20	500-50-0505	Jose	Rodriguez	7/16/1998	$150,000	M	Good
21	555-22-3333	Patricia	Rubin	7/25/2003	$45,000	F	Average
22	555-56-5555	Kenneth	Charles	6/18/1998	$40,000	M	Poor
23	612-99-1111	Melissa	Roberts	5/14/1984	$79,000	F	Good
24	625-62-6262	Holly	Holmes	6/15/1992	$55,000	F	Average
25	767-74-7373	William	Martin	8/26/2006	$23,000	M	Good
26	776-67-6666	David	Adamson	10/4/2002	$52,000	M	Poor
27	777-78-7777	Kelly	Marder	9/25/1997	$38,500	F	Average
28	925-45-7116	David	Whitehead	7/25/1980	$175,000	M	Good

Filter only the data you want

Filter data syntax

```
=FILTER( employee, salary<100000 )
```

dataset

condition

Sort data

ssn	firstname	lastname	hiredate	salary	gender	performance
333-66-1234	Marietta	Brown	3/7/2001	$18,500	F	Poor
000-02-2222	Sandy	Adams	1/15/2001	$19,500	F	Average
767-74-7373	William	Martin	8/26/2006	$23,000	M	Good
777-78-7777	Kelly	Marder	9/25/1997	$38,500	F	Average
555-56-5555	Kenneth	Charles	6/18/1998	$40,000	M	Poor
222-52-5555	Mary	Smith	1/1/2006	$42,500	F	Average
555-22-3333	Patricia	Rubin	7/25/2003	$45,000	F	Average
111-12-1111	James	Johnson	5/3/1996	$47,500	M	Good
333-34-3333	Emily	Manin	12/1/2000	$49,500	F	Average
776-67-6666	David	Adamson	10/4/2002	$52,000	M	Poor
109-87-6544	Harold	Foster	8/14/2005	$55,000	M	Good
625-62-6262	Holly	Holmes	6/15/1992	$55,000	F	Average
000-01-0000	Patricia	Milgrom	10/1/2004	$57,500	F	Average
432-19-8765	Paul	Bronson	11/20/2003	$58,000	M	Good
464-64-4466	David	Webster	1/29/1991	$58,500	M	Poor
333-43-4444	Frank	Smith	1/29/1991	$65,000	M	Good
335-55-5533	Holly	Jones	4/8/1986	$65,000	F	Good
109-87-6543	Emily	Wood	3/12/1997	$69,000	F	Average
245-67-8910	Sandy	Johanson	6/2/2005	$69,000	F	
444-45-4444	Vernon	Frank	4/10/1985	$75,000	M	Good
612-99-1111	Melissa	Roberts	5/14/1984	$79,000	F	Good
123-45-6789	Tracy	Coulter	2/14/1993	$100,000		Good
222-23-2222	Bill	Marlin	3/28/1977	$125,000	M	
500-50-0505	Jose	Rodriguez	7/16/1998	$150,000	M	Good
925-45-7116	David	Whitehead	7/25/1980	$175,000	M	Good

Sort data from low to high (or high to low)

Sort data syntax

```
=SORT( employee, column_index, ascending )
```

dataset — column to sort — low to high

QUERY

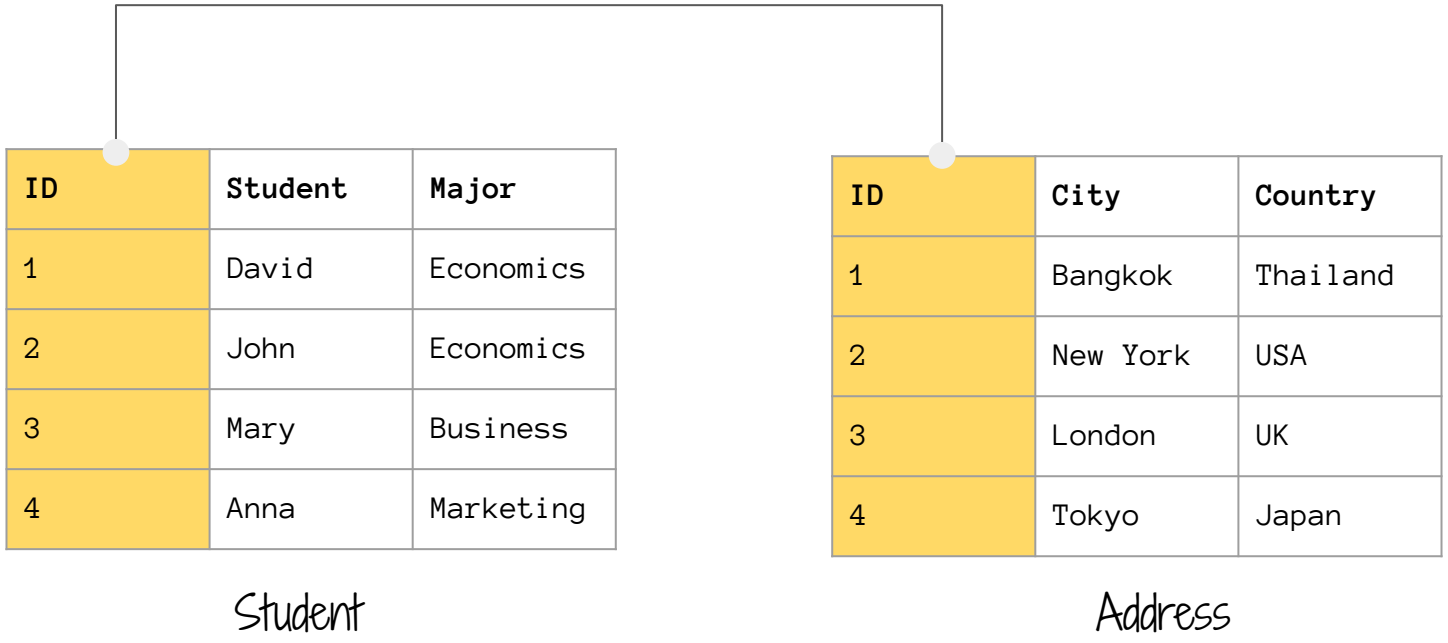
Select and Where clauses

```
=QUERY( employee, “select *” )
```

dataset

SQL like syntax

How to join two tables



ID	Student	Major
1	David	Economics
2	John	Economics
3	Mary	Business
4	Anna	Marketing

Student

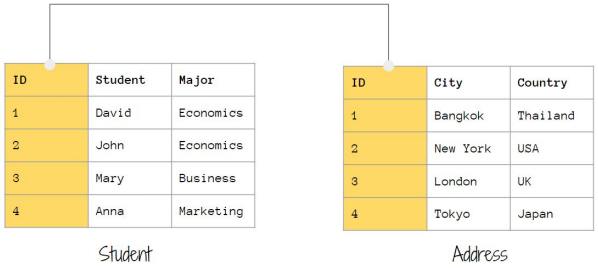
ID	City	Country
1	Bangkok	Thailand
2	New York	USA
3	London	UK
4	Tokyo	Japan

Address

Joined Table

ID	Student	Major	City	Country
1	David	Economics	Bangkok	Thailand
2	John	Economics	New York	USA
3	Mary	Business	London	UK
4	Anna	Marketing	Tokyo	Japan

We can join tables with VLOOKUP



ID	Student	Major
1	David	Economics
2	John	Economics
3	Mary	Business
4	Anna	Marketing

Student

ID	City	Country
1	Bangkok	Thailand
2	New York	USA
3	London	UK
4	Tokyo	Japan

Address

```
=VLOOKUP( id, address, columns, match )
```

id in student table

lookup table

columns you want

exact match

VLOOKUP

Employee Table

ssn	lastname	firstname	positionID	locationID
000-01-0000	Milgrom	Patricia	2	2
000-02-2222	Adams	Sandy	3	1
109-87-6543	Wood	Emily	2	5
109-87-6544	Foster	Harold	1	3
111-12-1111	Johnson	James	1	3
123-45-6789	Coulter	Tracy	2	1

Location Table (Lookup)

locationID	locationcity	address	state	zipcode	officephone
1	Atlanta	450 Peachtree	GA	30316	(404)333-5555
2	Boston	3 Commons Bl	MA	2190	(617)123-4444
3	Chicago	500 Loop High	IL	60620	(312)444-6666
4	Miami	210 Biscayne B	FL	33103	(305)787-9999
5	New York City	1650 Washingt	NY	15648	(518)256-3100
6	Denver	312 Mount Viev	CO	54657	(205)607-5289
7	Salt Lake City	316 S. State St	UT	84125	(801)459-6652
8	Los Angeles	1400 Main St	CA	94235	(705)639-0227

VLOOKUP

Employee Table

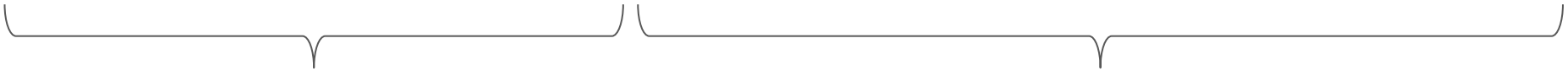
ssn	lastname	firstname	positionID	locationID
000-01-0000	Milgrom	Patricia	2	2
000-02-2222	Adams	Sandy	3	1
109-87-6543	Wood	Emily	2	5
109-87-6544	Foster	Harold	1	3
111-12-1111	Johnson	James	1	3
123-45-6789	Coulter	Tracy	2	1

Location Table (Lookup)

locationID	locationcity	address	state	zipcode	officephone
1	Atlanta	450 Peachtree	GA	30316	(404)333-5555
2	Boston	3 Commons Bl	MA	2190	(617)123-4444
3	Chicago	500 Loop High	IL	60620	(312)444-6666
4	Miami	210 Biscayne B	FL	33103	(305)787-9999
5	New York City	1650 Washingt	NY	15648	(518)256-3100
6	Denver	312 Mount Viev	CO	54657	(205)607-5289
7	Salt Lake City	316 S. State St	UT	84125	(801)459-6652
8	Los Angeles	1400 Main St	CA	94235	(705)639-0227

Final Table (Join Both Tables into One)

ssn	lastname	positionID	locationID	locationcity	address	state	zipcode	officephone
000-01-0000	Milgrom	2	2	Boston	3 Commons Blvd	MA	2190	(617)123-4444
000-02-2222	Adams	3	1	Atlanta	450 Peachtree Rd	GA	30316	(404)333-5555
109-87-6543	Wood	2	5	New York City	1650 Washington Blvd	NY	15648	(518)256-3100
109-87-6544	Foster	1	3	Chicago	500 Loop Highway	IL	60620	(312)444-6666
111-12-1111	Johnson	1	3	Chicago	500 Loop Highway	IL	60620	(312)444-6666
123-45-6789	Coulter	2	1	Atlanta	450 Peachtree Rd	GA	30316	(404)333-5555
222-23-2222	Marlin	2	4	Miami	210 Biscayne Blvd	FL	33103	(305)787-9999
222-52-5555	Smith	1	3	Chicago	500 Loop Highway	IL	60620	(312)444-6666
245-67-8910	Johanson	1	6	Denver	312 Mount View Dr	CO	54657	(205)607-5289
333-34-3333	Manin	1	2	Boston	3 Commons Blvd	MA	2190	(617)123-4444



Employee Table

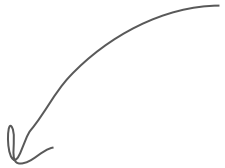
Location Table (Lookup)

You can search for pattern with Regular Expression

Google Sheets is the best! You can use it for free, 0$ cost.

Find G__ S__

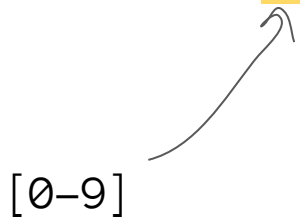
G[a-z]+ S[a-z]+



Google Sheets is the best! You can use it for free, 0$ cost.

Find 0-9 number

Google Sheets is the best! You can use it for free, 0$ cost.



Regular Expression Basics

`^A`	Ant, Amsterdam, America
`s$`	Toys, SNSDs, APPLEs
`c.t`	cat, cot, cet, cCt, c8t

Regular Expression Character Class

```
[ABC]  match A B or C
[A-Z]  match all capital letters
[A-z]  match all letters
[a-z]  match all lowercase letters
[0-9]  match digits
```

Regular Expression Quantifiers

```
*       match zero or more
+       match one or more
?       match zero or one
{5}     match exactly 5 characters
{3,5}   match min 3, max 5 characters
```

Regular expression - Wikipedia

More Examples :)

`[0-9]{5}`	match exactly 5 digits
`apples?`	apple, apples
`^[AB][0-9]{4}`	A1150, B2324, A3599

REGEXMATCH

```
=REGEXMATCH( employee_name, “^[PM]” )
```

text

starts with P or M

Extract Text

Text	Extracted
I love hamburger	hamburger
I love hotdog	hotdog
I love pizza	pizza
I love onion	onion
I love fried chicken	fried chicken

REGEXEXTRACT

=**REGEXEXTRACT**(text, “I love (food_you_want)”)

text

food you want in ( )

Project - ID Card Parser

We will extract information from this card

ID 3-5522-87666-87-2

Miss. Carry Anna

Date of Birth 18 Jan 1995

Address 967 Tokyo Japan 10880

Expired Date 25 Apr 2025

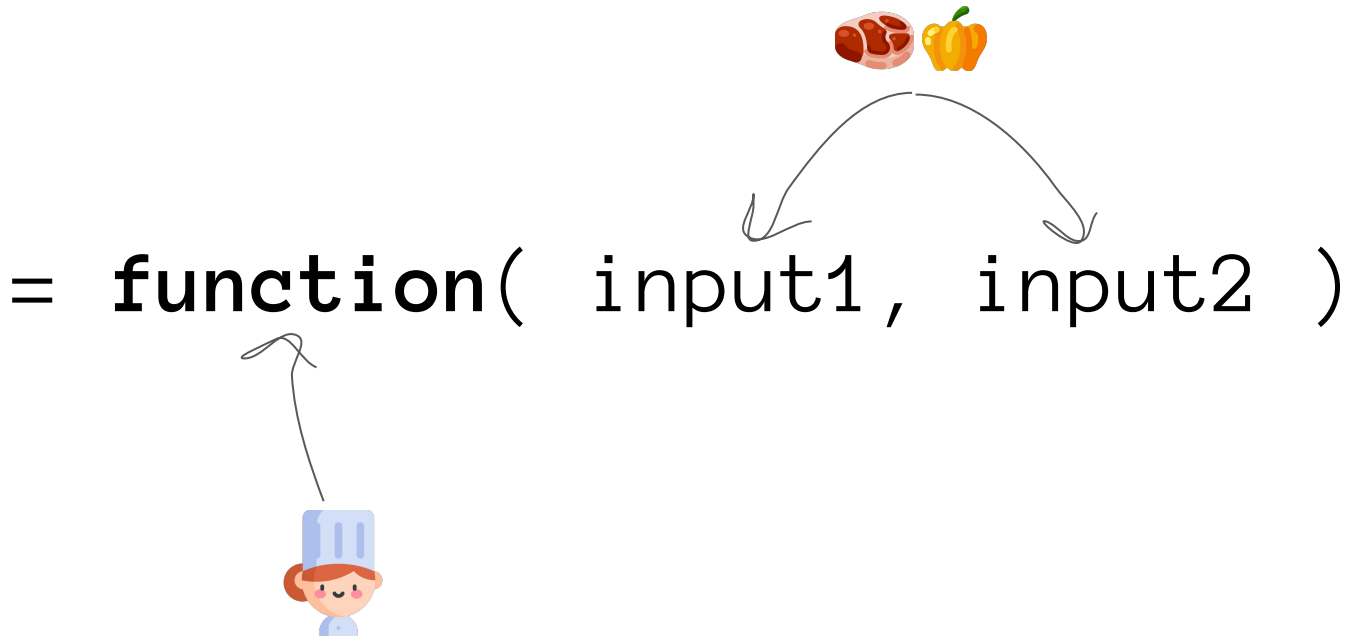
Pro Tip!

The more functions you know,
The more you can do

Bootcamp Live 01
Data Analytics with Google Sheets

Website: https://datarockie.com

A more technical view



The most popular function in sheets

```
= IF( score>=80 , “Pass”, “Fail”)
```