Introduction to Spreadsheets

What are formulas?

- Formulas are used to perform calculations on data in a spreadsheet
- Formulas always start with =
- Formulas consist of a function, like + or –, and arguments
- Arguments are what the function is applied to: either values (e.g., numbers) or references (defined on the next slide)
- An example of a simple formula is =1+2



What are references?

- A reference is a cell address, like A1 or A2 or B2 and so on
- An example of a formula involving references is =A1-B1



• If the value in A1 is 100, and the value in cell A2 is 1,

then this formula evaluates to 99

<u>F</u> ile	<u>E</u> dit <u>V</u> iew J	nsert F <u>o</u> rma	t <u>T</u> ools <u>D</u> ate	a <u>W</u> indow <u>H</u> elp
	- 🗷 - 🗔 🛛	🌫 🕑 🔒	🖴 🗣 I 🍣	ABS 😹 🖻 🖄
	Liberation	Sans 🔻 🛽	.0 T B	
SUMP	PRODUCT	- Fx 🗙	🛹 🖃	
	A	В	С	D
1	10	0	1 =A1·B1	
2				
3				

<u>F</u> ile	<u>E</u> dit <u>V</u> iew <u>I</u> r	nsert F <u>o</u> rmat	<u>T</u> ools <u>D</u> ata	<u>W</u> indow <u>H</u> elp
	- 🔁 - 🗖 🛛	a 🕑 🔒 🖁	🖹 🗣 I 🍄 🐴	🖌 l 😹 🖻 🛍
	Liberation S	Sans 🔻 10	• B I	<u>U</u> ∣≣ ≝ :
E15		• Jx 🗵		
	A	В	С	D
1	100	0 1	. 99	
2				
2				

More about formulas

- +, -, *, / are called binary functions, because they apply to two numbers!
- can also be a unary function, when it is the negative operator applied to just one number
- But some functions can be applied to an arbitrary number of arguments, like: sum, average, min, max, etc.

The syntax for formulas involving these functions is:

- Like other formulas, they always start with =
- = is followed by a function name, and parentheses
- Inside the parentheses is a data range, such as A1:A5





Example dataset: Revenue and Expenses of (a few) Ivy League Schools

- From financial year 2016
- Money is in millions of US dollars
- Source is from various schools' financial reports

	A	B	С	D	E	F	G	
1	School	President First Name	President Last Name	Endowment Support	Student Fees	Student Aid	Operating Expense	
2	Brown	Christina	Paxson	153.5	307.3	122.5	765.1	
3	Princeton	Christopher	Eisgruber	822.5	332.5	297.5	332.5	
4	Dartmouth	Philip	Hanlon	189.10	340.9	141.7	592.1	
-								

First, let's compute Brown's total revenue

- In column H create a new column and label it Total Revenue
- Click on cell H2, the first cell in this new column
- In the Formula bar type =
- Then click on the cell in row 2 of the column labeled money Endowment Support, type +, and then click on the cell in row 2 of the column labeled money Student Fees
- Hit enter and observe the total amount of money raised by Brown

	A	В	C	D	E	F	G	H
1 Sc	chool	President First Name	President Last Name	Endowment Support	Student Fees	Student Aid	Operating Expense	Total Revenue
2 Br	rown	Christina	Paxson	153.5	307.3	122.5	765.1	460.8
3 Pr	rinceton	Christopher	Eisgruber	822.5	332.5	297.5	332.5	
4 D	artmouth	Philip	Hanlon	189.10	340.9	141.7	592.1	

Next, let's compute all these schools' total revenues

- Click on the lower right corner of H2, and drag your mouse down to populate the next 3 cells.
- Click on any of those 3 cells, and observe the formula that appears in the formula bar for that cell.
- Spreadsheets are very "smart". By default, they assume relative references: they adjust formulas as necessary to reference the relative row or column.
- Absolute references are specified using dollar signs: E.g.,
 - \$A1 to make the roq absolute
 - A\$1 to make the column absolute
 - \$A\$1 to make both the row and the column absolute

Total Revenue

	A	В	C	D	E	F	G	Н
1	School	President First Name	President Last Name	Endowment Support	Student Fees	Student Aid	Operating Expense	Total Revenue
2	Brown	Christina	Paxson	153.5	307.3	122.5	765.1	460.8
3	Princeton	Christopher	Eisgruber	822.5	332.5	297.5	332.5	
4	Dartmouth	Philip	Hanlon	189.10	340.9	141.7	592.1	



	H	
ıg	Total Revenue	T S
55.1	460.8	
32.5	1155	
92.1	530	



Who had the highest total revenue?

- Click on the cell at the bottom of the Total Revenue column
- In the Formula bar write =MAX (
- Then click on first cell under Total Revenue, type :, and then click on the bottom cell under Total Revenue
- Then end the formula with), and hit Enter

Н	I		
	Total		r
School	Revenue		
Brown		460.8	
Princeton		1155	
Dartmouth		530	
	=MAX(I2:I4)		



Who had the lowest total revenue?

- Click on the cell at the bottom of the Total Revenue column
- In the Formula bar write =MIN (
- Then click on first cell under Total Revenue, type :, and then click on the bottom cell under Total Revenue
- Then end the formula with), and hit Enter

Н	1
School	Total Revenue
Brown	460.8
Princeton	1155
Dartmouth	530
	=MIN(12:14)

	Н		1
		Total	,
	School	Revenue	
(Brown	1	460.8
1	Princeton	1	1155
	Dartmouth	l I	530
			1155
			T

Now, let's compute Brown's total expenses

- In column I create a new column and label it Total Expenses
- Click on cell 12, the first cell in this new column
- In the Formula bar type =
- Then click on the cell in row 2 of the column labeled Student Aid, type +, and then click on the cell in row 2 of the column labeled Operating Expense
- Hit enter and observe the total amount of money spent by Brown

	A	B	С	D	E	F	G	H	1
1	School	President First Name	President Last Name	Endowment Support	Student Fees	Student Aid	Operating Expense	Total Revenue	Total Expenses ₇
2	Brown	Christina	Paxson	153.5	307.3	122.5	765.1	460.8	887.6
3	Princeton	Christopher	Eisgruber	822.5	332.5	297.5	332.5	1155	
4	Dartmouth	Philip	Hanlon	189.10	340.9	141.7	592.1	530	

Next, let's compute all these schools' total expenses

- Click on the lower right corner of 12, and drag your mouse down to populate the next 3 cells.
- Click on any of those 3 cells, and observe the formula that appears in the formula bar for that cell.
- Spreadsheets are very "smart". They adjust formulas as necessary to reference the appropriate row.

Total Expenses

	A	B	C	D	E	F	G	H	1
1	School	President First Name	President Last Name	Endowment Support	Student Fees	Student Aid	Operating Expense	Total Revenue	Total Expenses
2	Brown	Christina	Paxson	153.5	307.3	122.5	765.1	460.8	887.6
3	Princeton	Christopher	Eisgruber	822.5	332.5	297.5	332.5	1155	
4	Dartmouth	Philip	Hanlon	189.10	340.9	141.7	592.1	530	
		10.2							



I	
Total Expenses	5
887.6	
630	
733.8	
	 Total Expenses 887.6 630 733.8

Who had the highest total expenses?

- Click on the cell at the bottom of the Total Expenses column
- In the Formula bar write =MAX (
- Then click on first cell under Total Expenses, type :, and then click on the bottom cell under Total Expenses
- Then end the formula with), and hit Enter

Н	1	J	Н	1	J
School	Total Revenue	Total Expenses	School	Total Revenue	Total Expense
Brown	460.8	887.6	Brown	460.8	887.6
Princeton	1155	630	Princeton	1155	630
Dartmouth	530	733.8	Dartmouth	530	733.8
	460.8	=MAX (J2:J4)		460.8	887.6

How about finding the average amount of money spent?

- Click on the cell at the bottom of the Total Expenses column
- In the Formula bar write =AVERAGE (
- Then click on the top cell under Total Expenses, type :, and then click on the bottom cell under Total Expenses
- End the formula with), and hit Enter

=AVERAG	E(12:14)			
Н	I			
otal	Total			
evenue	Expenses			
460.8	887.6			
1155	630			
530	733.8			
	=AVERAGE(I2:I4)			

=AVERAG	E(12:14)		
Н		I	
otal	Total		1
evenue	Expenses		[]
460.8			887.6
1155			630
530			733.8
			750.47
			12

Now, let's subtract total expenses from total revenue to determine total profits

- Create a new column (in column J) and label it Total Profit
- Click on the first cell in this new column, J2
- In the Formula bar type =
- Then click on H2; then enter –; then click on I2
- Click on the lower right-hand corner of the cell you just populated, and drag it down to populate the 3 cells below it
- Click on any of those 3 cells, and observe the formula that appears in the formula bar for that cell

J2	I_{12} f_{*} $\Sigma = = H_{2}$									
	A	В	C	D	E	F	G	H		J
1	School	President First Name	President Last Name	Endowment Support	Student Fees	Student Aid	Operating Expense	Total Revenue	Total Expenses	Total Profit
2	Brown	Christina	Paxson	153.5	307.3	122.5	765.1	460.8	887.6	-426.8
3	Princeton	Christopher	Eisgruber	822.5	332.5	297.5	332.5	1155	630	525
4	Dartmouth	Philip	Hanlon	189.10	340.9	141.7	592.1	530	733.8	-203.8 ¹

=H3-I3					
;	H	1	J		
ing se	Total Revenue	Total Expenses	Total Profit		
765.1	460.8	887.6	-426.8		
332.5	1155	630	525		
592 . 1	530	733.8	-203.8		

Who had the lowest profits?



Some formulas manipulate text

- Len: returns the length of a string
- Concatenate: combines two strings into one
- Trim: removes duplicate spaces, and spaces at the start and end of a string
- Exact: tests if two supplied text strings are exactly the same and if so, returns true, else it returns false. (This function is case sensitive.)

Let's see an example of concatenate

- Click on column B
- From the drop-down menus: Insert -> Column left
- Name this new column President
- Click on cell B2
- In the Formula bar, write =CONCATENATE (C2, "", D2)
- As usual, drag the contents of B2 down to also populate the 3 cells below it

\checkmark f _* Σ = =CONCATENATE(C2, "", D2)				
В	С	D		
President	President First Name	President Last Name		
Christina Paxson	Christina	Paxson		
Christopher Eisgruber	Christopher	Eisgruber		
Philip Hanlon	Philip	Hanlon		

\checkmark f _* Σ = = =CONCATENA	ATE(C3, " ", D3)		
В	С	D	
President	President First Name	President Last Name	
Christina Paxson	Christina	Paxson	
Christopher Eisgruber	Christopher	Eisgruber	
Philip Hanlon	Philip	Hanlon	

Summary

- Spreadsheets are incredibly powerful tools, and easy to use!
- Using formulas, we can manipulate data to compute summary statistics
 MIN, MAX, AVERAGE, etc.
- We can also visualize data in spreadsheets, so we can more easily identify patterns, and uncover stories that are hidden in our data.
- However, we will also discover that they become unwieldy as our analyses become more complex.