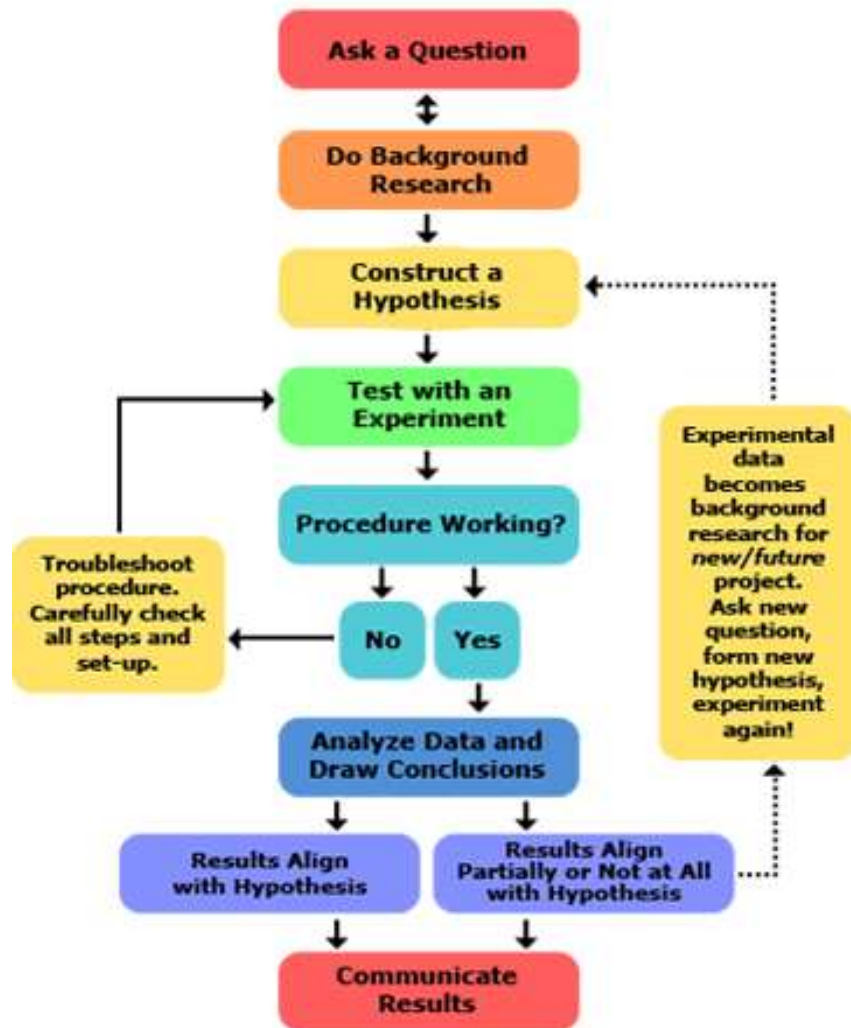


Doing Science : from Start to Finish (day 5)

Dr. Félix E. Rivera-Mariani

<https://github.com/friveramariani/DoingScienceWorkshops>

The Scientific Method

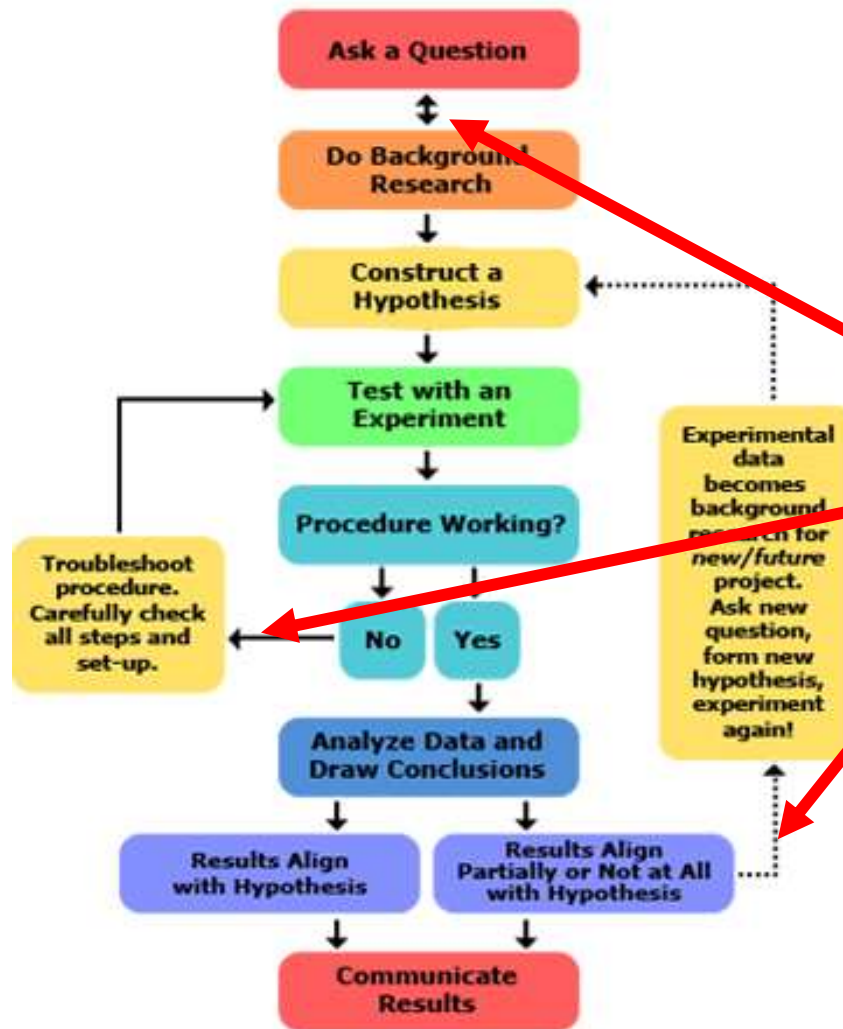


The Scientific Method: Epicycles



- Two directions between question and background research

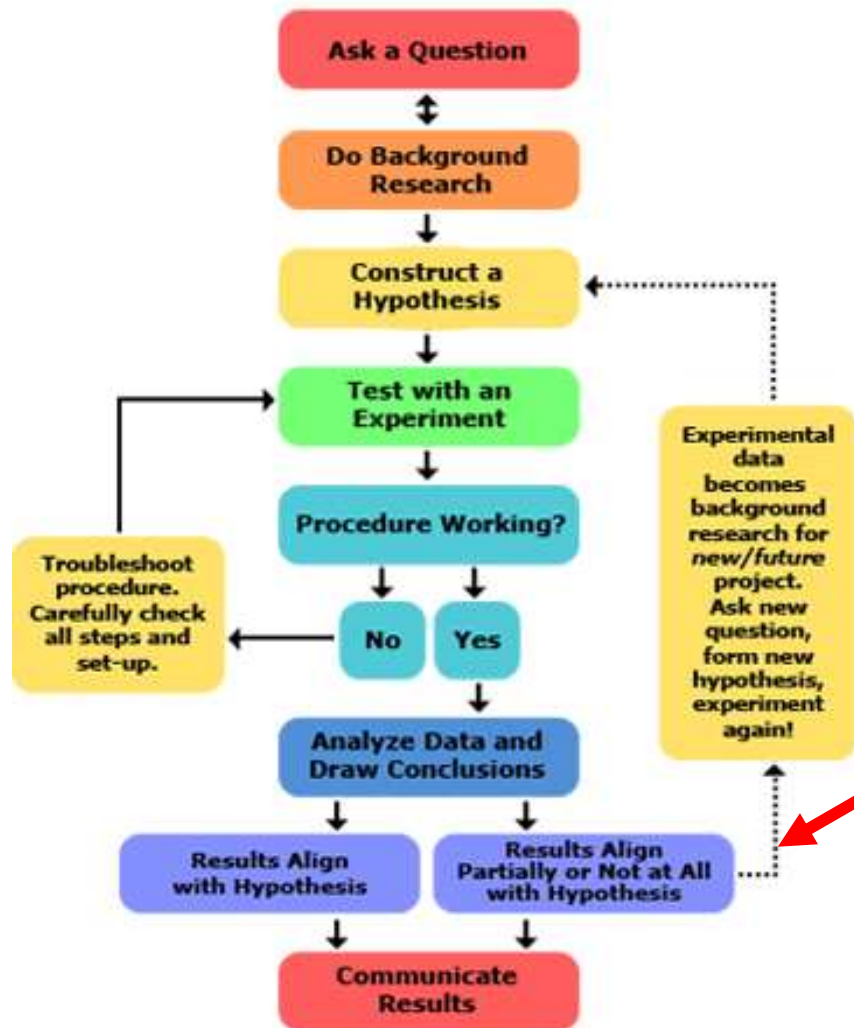
The Scientific Method: Epicycles



- Two directions between question and background research

Epicycles with the experiment

The Scientific Method: Epicycles

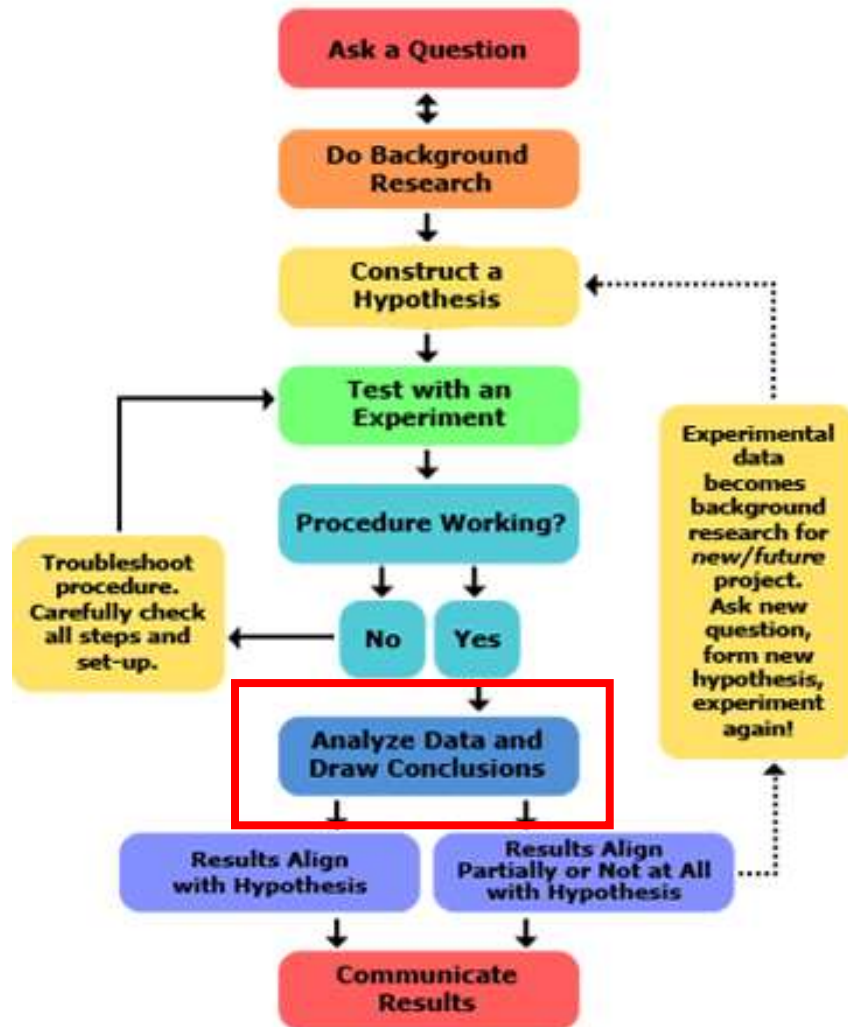


- Two directions between question and background research

- Epicycles with the experiment

Epicycle between results and hypothesis

Today's goal: Data



Data

The Scientific Method: Data Formats

During data analysis (not necessarily data collection)...

a) Columns = variables

b) Rows = observations

Analyze Data and
Draw Conclusions

Incorrect

hyp1	hyp2
685	155
216	183
137	92
226	147
101	124
127	124
76	140
117	143
	78
	120

Correct

hypothesis	letters
hyp1	685
hyp1	216
hyp1	137
hyp1	226
hyp1	101
hyp1	127
hyp1	76
hyp1	117
hyp2	155
hyp2	183
hyp2	92
hyp2	147
hyp2	124
hyp2	124
hyp2	140
hyp2	143
hyp2	78
hyp2	120

The Scientific Method: Data Formats

YearMonth	Tmax	Tmin	Tavg	WetBulb	PrecipTotal	Sunrise	Sunset	RH
1/1/2011	81	69	75	68	0.22	06:56	17:59	71.58621
1/2/2011	79	71	75	69	0.38	06:56	#N/A	77.58824
1/3/2011	80	70	75	67	0.18	06:57	18:00	67.45833
1/4/2011	79	70	75	69	0.08	06:57	18:01	70.23077
1/5/2011	80	71	76	71	0.86	06:57	18:01	84.42857
1/6/2011	82	71	77	72	0	06:58	18:02	78.83333
1/7/2011	86	70	78	73	0	06:58	18:03	76.66667
1/8/2011	89	73	81	73	0	06:58	18:03	74.41667
1/9/2011	84	72	78	73	0.01	06:58	18:04	80.79167

National Oceanic Atmospheric Administration meteorological variables for San Juan, PR
(www.noaa.gov)

Types of Data (Variables): **two main types**

Types of Data (Variables): two main types

Independent

Dependent

Types of Data (Variables): two main types

YearMonth	Tmax	Tmin	Tavg	WetBulb	PrecipTotal	Sunrise	Sunset	RH
1/1/2011	81	69	75	68	0.22	06:56	17:59	71.58621
1/2/2011	79	71	75	69	0.38	06:56	#N/A	77.58824
1/3/2011	80	70	75	67	0.18	06:57	18:00	67.45833
1/4/2011	79	70	75	69	0.08	06:57	18:01	70.23077
1/5/2011	80	71	76	71	0.86	06:57	18:01	84.42857
1/6/2011	82	71	77	72	0	06:58	18:02	78.83333
1/7/2011	86	70	78	73	0	06:58	18:03	76.66667
1/8/2011	89	73	81	73	0	06:58	18:03	74.41667
1/9/2011	84	72	78	73	0.01	06:58	18:04	80.79167

How many variables are dependent in this dataset?

Types of Data (Variables): two main types


Independent

- Already known
- Affects the independent variable

Dependent

- Changes as the independent changes

The Scientific Method: two main types



Already known

YearMonth	Tmax	Tmin	Tavg	WetBulb	PrecipTotal	Sunrise	Sunset	RH
1/1/2011	81	69	75	68	0.22	06:56	17:59	71.58621
1/2/2011	79	71	75	69	0.38	06:56	#N/A	77.58824
1/3/2011	80	70	75	67	0.18	06:57	18:00	67.45833
1/4/2011	79	70	75	69	0.08	06:57	18:01	70.23077
1/5/2011	80	71	76	71	0.86	06:57	18:01	84.42857
1/6/2011	82	71	77	72	0	06:58	18:02	78.83333
1/7/2011	86	70	78	73	0	06:58	18:03	76.66667
1/8/2011	89	73	81	73	0	06:58	18:03	74.41667
1/9/2011	84	72	78	73	0.01	06:58	18:04	80.79167

The Scientific Method: two main types

Already known

YearMonth	Tmax	Tmin	Tavg	WetBulb	PrecipTotal	Sunrise	Sunset	RH
1/1/2011	81	69	75	68	0.22	06:56	17:59	71.58621
1/2/2011	79	71	75	69	0.38	06:56	#N/A	77.58824
1/3/2011	80	70	75	67	0.18	06:57	18:00	67.45833
1/4/2011	79	70	75	69	0.08	06:57	18:01	70.23077
1/5/2011	80	71	76	71	0.86	06:57	18:01	84.42857
1/6/2011	82	71	77	72	0	06:58	18:02	78.83333
1/7/2011	86	70	78	73	0	06:58	18:03	76.66667
1/8/2011	89	73	81	73	0	06:58	18:03	74.41667
1/9/2011	84	72	78	73	0.01	06:58	18:04	80.79167

Not known unless we measure it

How many variable are independent?

How many variables are independent in this dataset?

name	Quiz1	Quiz2	Quiz3	Quiz4
Edgar E Alv	0	4	6	1.8
Karen Day	9	5	7	6
Dylan Cast	7	11	10	8.8
Judette Ce	3	5	6	6
Leidy Cruz	9.5	0	5	2
Brian Davi	10	7	8	5.2
Utisemine	2	3	7	6
Jaqueline F	12	8	9	7.5
Kassandra	10	5	9	8.3
Yailin Ferra	0	6	3	10.7
Marcos Fig	8	5	5	7.8
Christina F	10	12	12	12
Byran J Fue	0	7	0	4.7
Jessie Gon	5	8	7	2.8
Dianny Inig	8	5	6	3
Carlos E Le	5.5	3	3	7.2
Jose Rober	5.5	7	9	10.7

How many variable are independent?

How many variables are independent in this dataset?

Names and quizzes (Quiz 1, Quiz 2, Quiz 3)

name	Quiz1	Quiz2	Quiz3	Quiz4
Edgar E Alv	0	4	6	1.8
Karen Day	9	5	7	6
Dylan Cast	7	11	10	8.8
Judette Ce	3	5	6	6
Leidy Cruz	9.5	0	5	2
Brian Davi	10	7	8	5.2
Utisemine	2	3	7	6
Jaqueline F	12	8	9	7.5
Kassandra	10	5	9	8.3
Yailin Ferra	0	6	3	10.7
Marcos Fig	8	5	5	7.8
Christina F	10	12	12	12
Byran J Fue	0	7	0	4.7
Jessie Gon	5	8	7	2.8
Dianny Inig	8	5	6	3
Carlos E Le	5.5	3	3	7.2
Jose Rober	5.5	7	9	10.7

Types of Data (Variables): **within dependent and independent**

Types of Data (Variables): within dependent and independent

Continuous

Non-continuous

Type of variables

What type of non-continuous variable are names and quizzes (Quiz 1, Quiz 2, Quiz 3)?

name	Quiz1	Quiz2	Quiz3	Quiz4
Edgar E Alv	0	4	6	1.8
Karen Day	9	5	7	6
Dylan Cast	7	11	10	8.8
Judette Ce	3	5	6	6
Leidy Cruz	9.5	0	5	2
Brian Davi	10	7	8	5.2
Utisemine	2	3	7	6
Jaqueline F	12	8	9	7.5
Kassandra	10	5	9	8.3
Yailin Ferr	0	6	3	10.7
Marcos Fig	8	5	5	7.8
Christina F	10	12	12	12
Byran J Fue	0	7	0	4.7
Jessie Gon	5	8	7	2.8
Dianny Inig	8	5	6	3
Carlos E Le	5.5	3	3	7.2
Jose Rober	5.5	7	9	10.7

Types of Data (Variables): within dependent and independent

Continuous

- numerical

Non-continuous

Types of Data (Variables): within dependent and independent

Continuous

- numerical

Non-continuous

- Categorical

Types of Data (Variables): within dependent and independent

Continuous

- numerical

Non-continuous

- Categorical
 - Binary (dichotomous)

Types of Data (Variables): within dependent and independent

Continuous

- numerical

Non-continuous

- Categorical
 - Binary (dichotomous)
 - Ordinal (has order)
 - Nominal (not ordered)

Examples of non-continuous variables

Ordinal

Nominal