

# Example #1

What is/are the candidate key(s) in the following table? What is the strongest normal form respected by the table?

<b>employee_id</b>	<b>employee_name</b>	<b>salary</b>	<b>dept_name</b>	<b>dept_head</b>	<b>dept_location</b>
1	Sam	65000	IT	Dale	London
2	Pam	120000	HR	Taylor	Sydney
3	Simon	62000	IT	Dale	London
4	Mary	90500	HR	Taylor	Sydney
5	Todd	62910	IT	Dale	London

# Example #1

Candidate keys are: {**employee\_id**} and {**employee\_name**}

The table is in **2NF**. All of the features outside the candidate keys depend on the entire key, but there is a transitive dependency from dept\_name => dept\_head and dept\_name => dept\_location

employee_id	employee_name	salary	dept_name	dept_head	dept_location
1	Sam	65000	IT	Dale	London
2	Pam	120000	HR	Taylor	Sydney
3	Simon	62000	IT	Dale	London
4	Mary	90500	HR	Taylor	Sydney
5	Todd	62910	IT	Dale	London

# Example #2

What is/are the candidate key(s) in the following table? What is the strongest normal form respected by the table?

patient_id	surgeon_id	surgery	drug_admin
1111	145	Gallstones removal	Penicillin
1111	311	Kidney stones removal	none
1234	243	Ete cataract	Tetracycline
1234	467	Thrombosis	none
2345	189	Open Heart Surgery	Cephalosporin
4876	145	Gallstones removal	Demicillin
5123	145	Gallstones removal	none

# Example #2

Candidate key is: {**patient\_id**, **surgeon\_id**}

The table is in **3NF**. All of the features outside the non-candidate keys depend on the entire key, and none of them are functionally dependent on other features.

<b>patient_id</b>	<b>surgeon_id</b>	<b>surgery</b>	<b>drug_admin</b>
1111	145	Gallstones removal	Penicillin
1111	311	Kidney stones removal	none
1234	243	Eye cataract	Tetracycline
1234	467	Thrombosis	none
2345	189	Open Heart Surgery	Cephalosporin
4876	145	Gallstones removal	Demicillin
5123	145	Gallstones removal	none

# Example #3

What is/are the candidate key(s) in the following table? What is the strongest normal form respected by the table?

<b>professor</b>	<b>course_codes</b>
Prof. Arnold	(DSST289, DSST389)
Prof. X	(DSST289, DSST389)
Prof. Orr	(DSST189, MATH395)
Prof. Kvam	(DSST189, MATH329, MATH330)

# Example #3

Candidate key is: {**professor**}

The table is in **0NF**. There is more than one value in a cell, which breaks even the 1NF rules. Note that the rules are cumulative; it does not matter that it results in the 2NF and 3NF rules.

<b>professor</b>	<b>course_codes</b>
Prof. Arnold	(DSST289, DSST389)
Prof. X	(DSST289, DSST389)
Prof. Orr	(DSST189, MATH395)
Prof. Kvam	(DSST189, MATH329, MATH330)

# Example #4

What is/are the candidate key(s) in the following table? What is the strongest normal form respected by the table?

*Note: This is the way to fix the 1NF issue with the Example 3 table.*

professor	course_codes
Prof. Arnold	DSST289
Prof. Arnold	DSST389
Prof. X	DSST289
Prof. X	DSST389
Prof. Orr	DSST189
Prof. Orr	MATH395
Prof. Kvam	DSST189
Prof. Kvam	MATH329
Prof. Kvam	MATH330

# Example #4

Candidate key is: {**professor**, **course\_codes**}

The table is in **3NF**. Note: when the candidate key consists of all the columns, 1NF always implies 3NF.

<b>professor</b>	<b>course_codes</b>
Prof. Arnold	DSST289
Prof. Arnold	DSST389
Prof. X	DSST289
Prof. X	DSST389
Prof. Orr	DSST189
Prof. Orr	MATH395
Prof. Kvam	DSST189
Prof. Kvam	MATH329
Prof. Kvam	MATH330



# Example #5

What is/are the candidate key(s) in the following table? What is the strongest normal form respected by the table?

city	country	continent	is_large
Paris	France	Europe	TRUE
Paris	United States	North America	FALSE
London	United Kingdom	Europe	TRUE
London	Canada	North America	TRUE
London	United States	North America	TRUE
Toronto	Canada	North America	TRUE
Athens	Greece	Europe	TRUE

# Example #5

Candidate key is: {city, country}

The table is in **1NF** because continent depends only on country, a proper subset of the first candidate key and country depends on a proper subset of the second candidate key.

city	country	continent	is_large
Paris	France	Europe	TRUE
Paris	United States	North America	FALSE
London	United Kingdom	Europe	TRUE
London	Canada	North America	TRUE
London	United States	North America	TRUE
Toronto	Canada	North America	TRUE
Athens	Greece	Europe	TRUE

# Data Feminism

List the four principles of *Data Feminism* as described on the class slides. A single sentence for each is sufficient.

# Data Feminism

## DATA IS A STRUCTURE OF POWER

1. Align work with your values to affect positive change in the world.

## DATA IS NEVER NEUTRAL OR OBJECTIVE

2. Reflect, align with values, and document all the decisions made when creating data.
3. Integrate a diverse range of sources; combine qualitative + quantitative analyses.

## DATA SCIENCE REQUIRES MANY PARTICIPANTS

4. Surface all the people that contribute to our work as data scientists.