Syracuse University
School of Information Studies

Data Normalization

Agenda

- What is data normalization?
- Understanding data anomalies and inconsistencies
- Defining the normal forms
- Understanding and identifying the normal forms
- Normalizing the table design and transitioning to other normal forms

Syracuse University
School of Information Studies

Data Normalization

The End

Syracuse University
School of Information Studies

What Is Normalization?

What Is Data Normalization?

- Data normalization is a process that seeks to improve the design of a relational database.
- The normalization process consists of a series of ordered tests over the data in each table.
- When a given test fails, there is a formal process we apply to improve the design.
- This involves redesigning the tables, then moving the data into the new versions of the tables.

Syracuse University
School of Information Studies

What Is Normalization?

The End

Syracuse University
School of Information Studies

Case Study

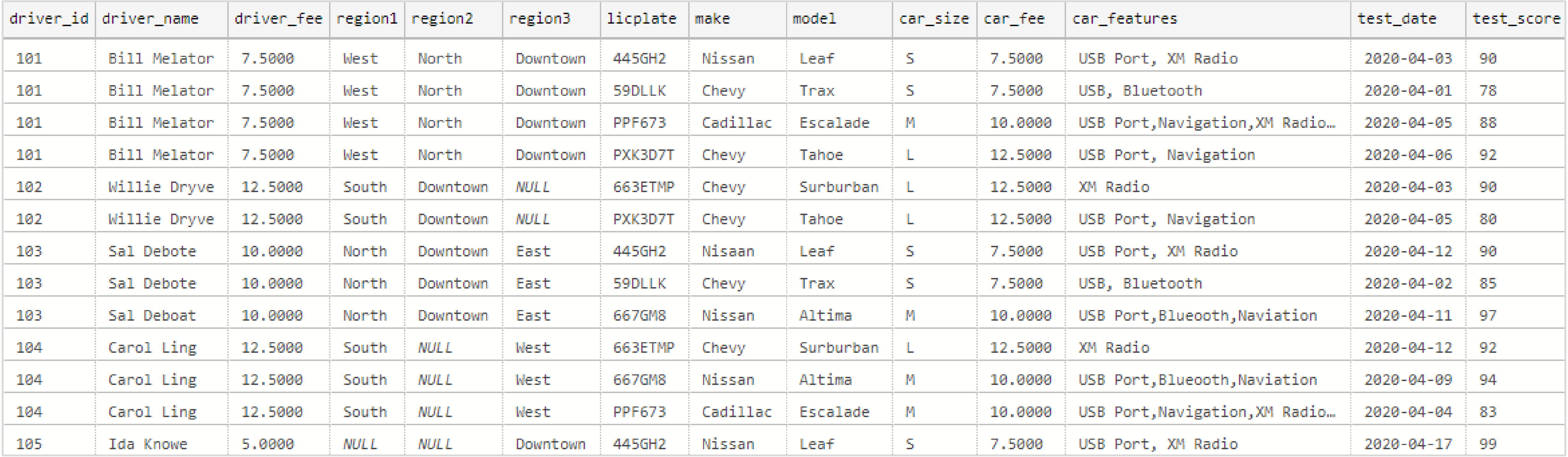
Cuse Rides

Case Study: Cuse Rides

- Community ride share program with fixed rates! (fictitious)
- Drivers: people who drive others needing a ride share
- Vehicles: fleet vehicles used to drive people around
- Each driver must be certified to drive that vehicle by scoring better than 75/100 on the road test exam
- Pricing model is based on driver and vehicle size, e.g., Bob driving large cars

Case Study: Cuse Rides (cont.)

- Spreadsheet database: one big table—poor design



driver_id	driver_name	driver_fee	region1	region2	region3	licplate	make	model	car_size	car_fee	car_features	test_date	test_score
101	Bill Melator	7.5000	West	North	Downtown	445GH2	Nissan	Leaf	S	7.5000	USB Port, XM Radio	2020-04-03	90
101	Bill Melator	7.5000	West	North	Downtown	59DLLK	Chevy	Trax	S	7.5000	USB, Bluetooth	2020-04-01	78
101	Bill Melator	7.5000	West	North	Downtown	PPF673	Cadillac	Escalade	M	10.0000	USB Port,Navigation,XM Radio…	2020-04-05	88
101	Bill Melator	7.5000	West	North	Downtown	PXK3D7T	Chevy	Tahoe	L	12.5000	USB Port, Navigation	2020-04-06	92
102	Willie Dryve	12.5000	South	Downtown	*NULL*	663ETMP	Chevy	Surburban	L	12.5000	XM Radio	2020-04-03	90
102	Willie Dryve	12.5000	South	Downtown	*NULL*	PXK3D7T	Chevy	Tahoe	L	12.5000	USB Port, Navigation	2020-04-05	80
103	Sal Debote	10.0000	North	Downtown	East	445GH2	Nisaan	Leaf	S	7.5000	USB Port, XM Radio	2020-04-12	90
103	Sal Debote	10.0000	North	Downtown	East	59DLLK	Chevy	Trax	S	7.5000	USB, Bluetooth	2020-04-02	85
103	Sal Deboat	10.0000	North	Downtown	East	667GM8	Nissan	Altima	M	10.0000	USB Port,Blueooth,Naviation	2020-04-11	97
104	Carol Ling	12.5000	South	*NULL*	West	663ETMP	Chevy	Surburban	L	12.5000	XM Radio	2020-04-12	92
104	Carol Ling	12.5000	South	*NULL*	West	667GM8	Nissan	Altima	M	10.0000	USB Port,Blueooth,Naviation	2020-04-09	94
104	Carol Ling	12.5000	South	*NULL*	West	PPF673	Cadillac	Escalade	M	10.0000	USB Port,Navigation,XM Radio…	2020-04-04	83
105	Ida Knowe	5.0000	*NULL*	*NULL*	Downtown	445GH2	Nissan	Leaf	S	7.5000	USB Port, XM Radio	2020-04-17	99

Syracuse University
School of Information Studies

Case Study

The End

Syracuse University
School of Information Studies

Data Inconsistencies

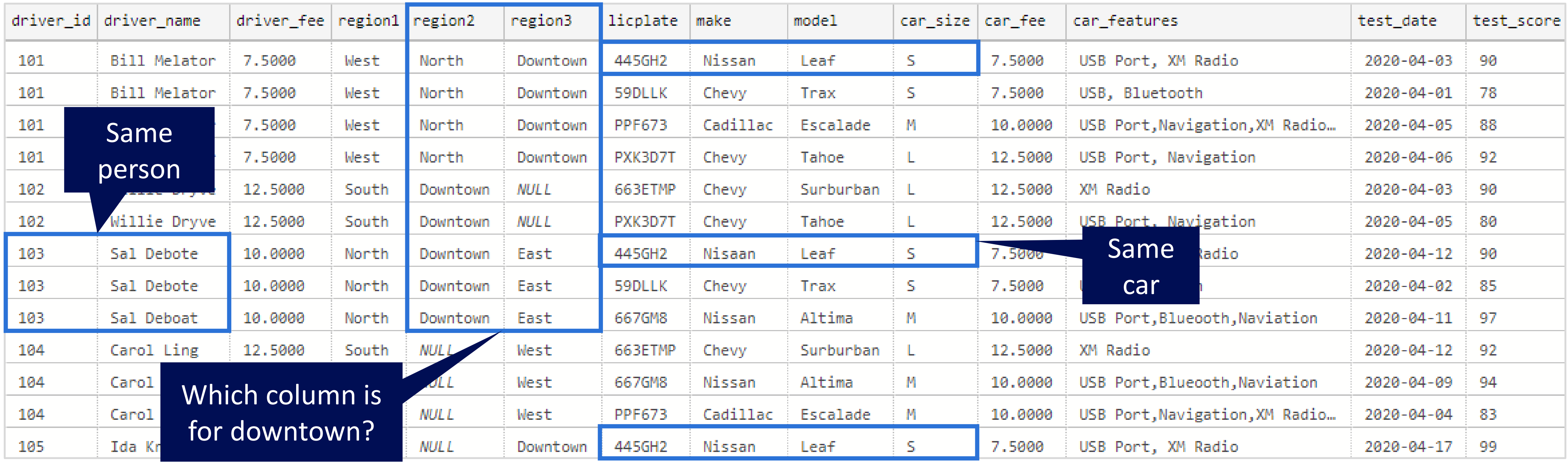
How Does Data Normalization Improve Database Design?

The process of normalization will identify hidden dependencies in the data that might have been overlooked throughout the design process.

These dependencies reveal themselves through data inconsistencies and data anomalies in the data within our tables.

Example: Data Inconsistencies

Multiple representations of the same data cause it to become inconsistent.



driver_id	driver_name	driver_fee	region1	region2	region3	licplate	make	model	car_size	car_fee	car_features	test_date	test_score
101	Bill Melator	7.5000	West	North	Downtown	445GH2	Nissan	Leaf	S	7.5000	USB Port, XM Radio	2020-04-03	90
101	Bill Melator	7.5000	West	North	Downtown	59DLLK	Chevy	Trax	S	7.5000	USB, Bluetooth	2020-04-01	78
101	[illegible]	7.5000	West	North	Downtown	PPF673	Cadillac	Escalade	M	10.0000	USB Port,Navigation,XM Radio…	2020-04-05	88
101	[illegible]	7.5000	West	North	Downtown	PXK3D7T	Chevy	Tahoe	L	12.5000	USB Port, Navigation	2020-04-06	92
102	[illegible]	12.5000	South	Downtown	*NULL*	663ETMP	Chevy	Surburban	L	12.5000	XM Radio	2020-04-03	90
102	Willie Dryve	12.5000	South	Downtown	*NULL*	PXK3D7T	Chevy	Tahoe	L	12.5000	USB Port, Navigation	2020-04-05	80
103	Sal Debote	10.0000	North	Downtown	East	445GH2	Nisaan	Leaf	S	7.5000	[illegible] Radio	2020-04-12	90
103	Sal Debote	10.0000	North	Downtown	East	59DLLK	Chevy	Trax	S	7.5000	[illegible]	2020-04-02	85
103	Sal Deboat	10.0000	North	Downtown	East	667GM8	Nissan	Altima	M	10.0000	USB Port,Blueooth,Naviation	2020-04-11	97
104	Carol Ling	12.5000	South	*NULL*	West	663ETMP	Chevy	Surburban	L	12.5000	XM Radio	2020-04-12	92
104	Carol [illegible]	[illegible]	[illegible]	*NULL*	West	667GM8	Nissan	Altima	M	10.0000	USB Port,Blueooth,Naviation	2020-04-09	94
104	Carol [illegible]	[illegible]	[illegible]	*NULL*	West	PPF673	Cadillac	Escalade	M	10.0000	USB Port,Navigation,XM Radio…	2020-04-04	83
105	Ida Kr[illegible]	[illegible]	[illegible]	*NULL*	Downtown	445GH2	Nissan	Leaf	S	7.5000	USB Port, XM Radio	2020-04-17	99

Same person

Same car

Which column is for downtown?

Syracuse University
School of Information Studies

Data Inconsistencies

The End

Syracuse University
School of Information Studies

Data Anomalies

Data Anomalies Create Data Inconsistencies

Update anomaly

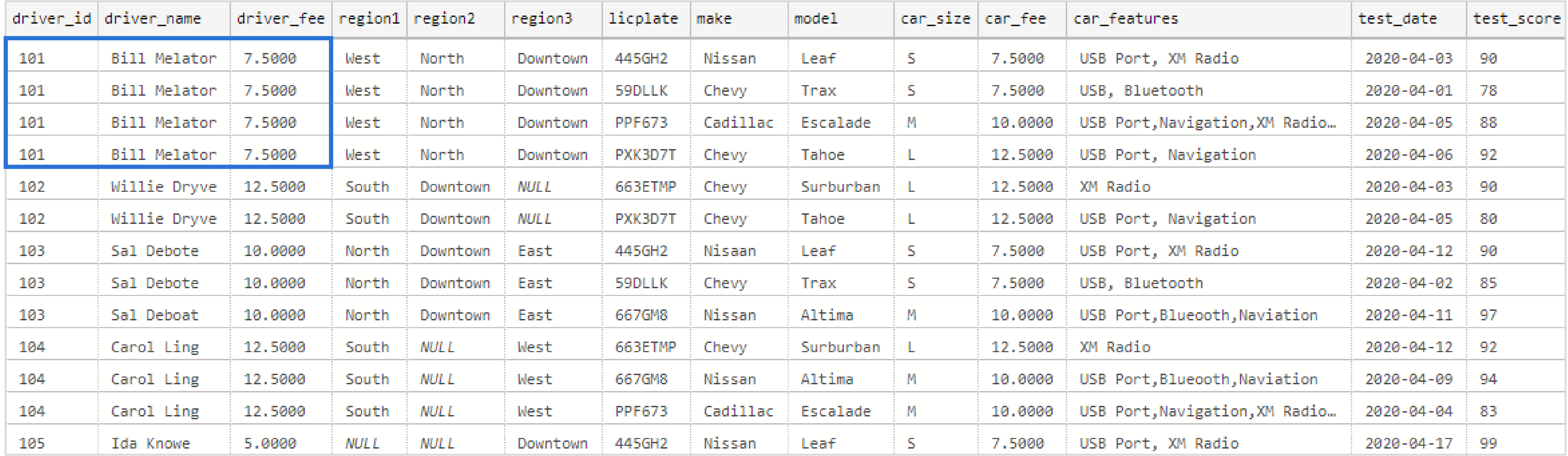
Delete anomaly

Insert anomaly

Table is not DRY!

Update Anomaly Example

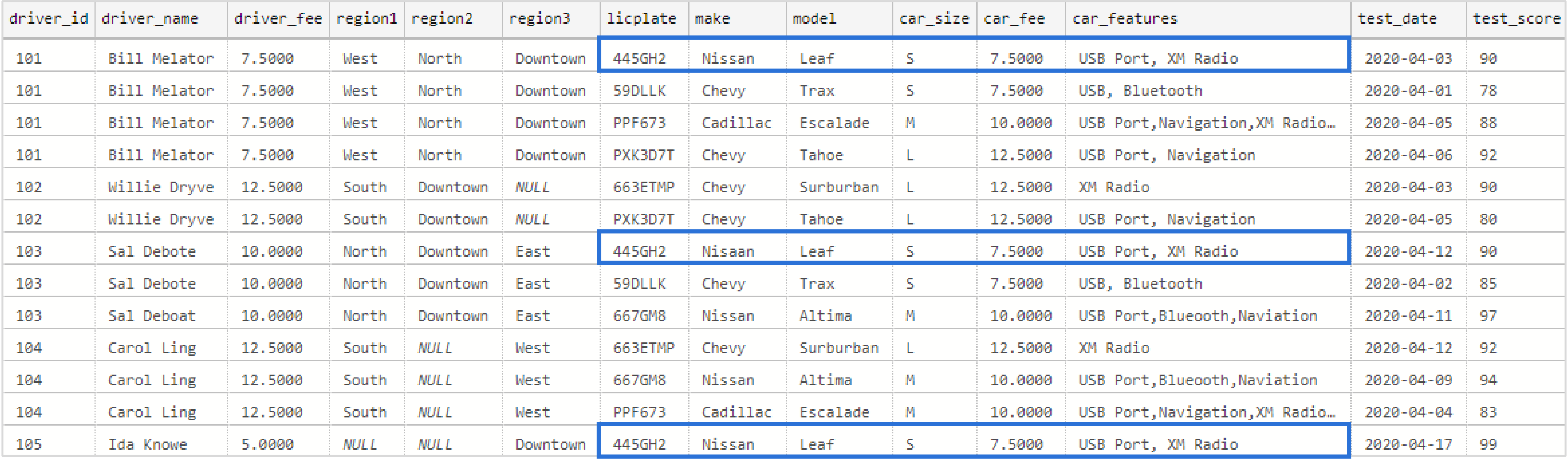
- Let’s change Bill Melator’s fee to 8.5.
- This requires a multiple-row update!



driver_id	driver_name	driver_fee	region1	region2	region3	licplate	make	model	car_size	car_fee	car_features	test_date	test_score
101	Bill Melator	7.5000	West	North	Downtown	445GH2	Nissan	Leaf	S	7.5000	USB Port, XM Radio	2020-04-03	90
101	Bill Melator	7.5000	West	North	Downtown	59DLLK	Chevy	Trax	S	7.5000	USB, Bluetooth	2020-04-01	78
101	Bill Melator	7.5000	West	North	Downtown	PPF673	Cadillac	Escalade	M	10.0000	USB Port,Navigation,XM Radio…	2020-04-05	88
101	Bill Melator	7.5000	West	North	Downtown	PXK3D7T	Chevy	Tahoe	L	12.5000	USB Port, Navigation	2020-04-06	92
102	Willie Dryve	12.5000	South	Downtown	*NULL*	663ETMP	Chevy	Surburban	L	12.5000	XM Radio	2020-04-03	90
102	Willie Dryve	12.5000	South	Downtown	*NULL*	PXK3D7T	Chevy	Tahoe	L	12.5000	USB Port, Navigation	2020-04-05	80
103	Sal Debote	10.0000	North	Downtown	East	445GH2	Nisaan	Leaf	S	7.5000	USB Port, XM Radio	2020-04-12	90
103	Sal Debote	10.0000	North	Downtown	East	59DLLK	Chevy	Trax	S	7.5000	USB, Bluetooth	2020-04-02	85
103	Sal Deboat	10.0000	North	Downtown	East	667GM8	Nissan	Altima	M	10.0000	USB Port,Blueooth,Naviation	2020-04-11	97
104	Carol Ling	12.5000	South	*NULL*	West	663ETMP	Chevy	Surburban	L	12.5000	XM Radio	2020-04-12	92
104	Carol Ling	12.5000	South	*NULL*	West	667GM8	Nissan	Altima	M	10.0000	USB Port,Blueooth,Naviation	2020-04-09	94
104	Carol Ling	12.5000	South	*NULL*	West	PPF673	Cadillac	Escalade	M	10.0000	USB Port,Navigation,XM Radio…	2020-04-04	83
105	Ida Knowe	5.0000	*NULL*	*NULL*	Downtown	445GH2	Nissan	Leaf	S	7.5000	USB Port, XM Radio	2020-04-17	99

Delete Anomaly Example

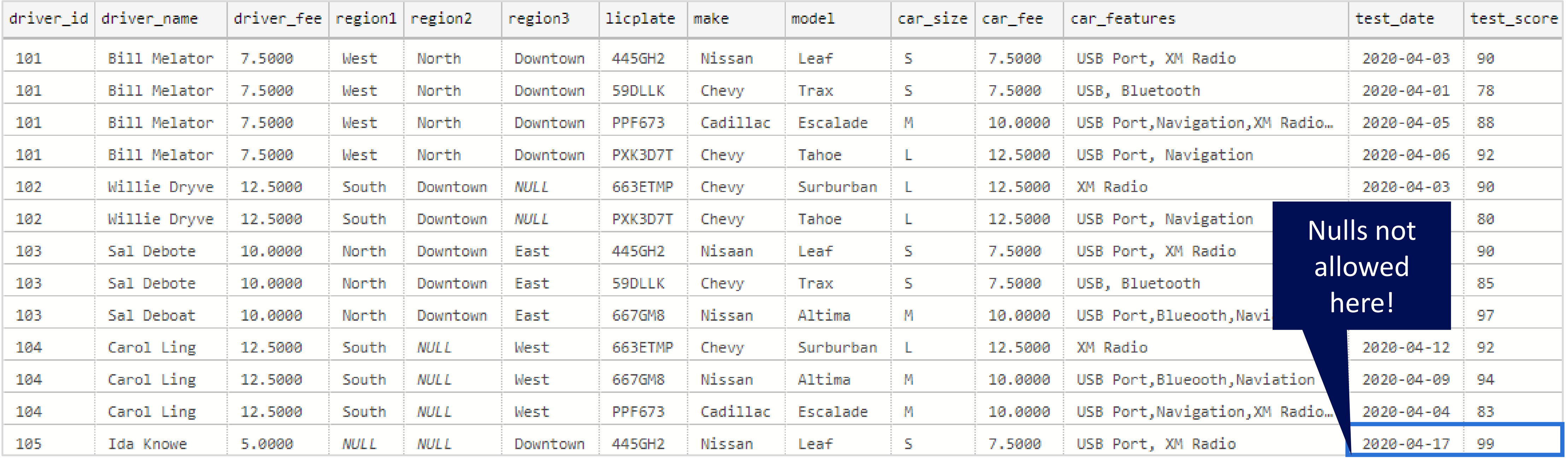
How do we delete the Nissan Leaf? A delete is an SQL UPDATE?



driver_id	driver_name	driver_fee	region1	region2	region3	licplate	make	model	car_size	car_fee	car_features	test_date	test_score
101	Bill Melator	7.5000	West	North	Downtown	445GH2	Nissan	Leaf	S	7.5000	USB Port, XM Radio	2020-04-03	90
101	Bill Melator	7.5000	West	North	Downtown	59DLLK	Chevy	Trax	S	7.5000	USB, Bluetooth	2020-04-01	78
101	Bill Melator	7.5000	West	North	Downtown	PPF673	Cadillac	Escalade	M	10.0000	USB Port,Navigation,XM Radio…	2020-04-05	88
101	Bill Melator	7.5000	West	North	Downtown	PXK3D7T	Chevy	Tahoe	L	12.5000	USB Port, Navigation	2020-04-06	92
102	Willie Dryve	12.5000	South	Downtown	*NULL*	663ETMP	Chevy	Surburban	L	12.5000	XM Radio	2020-04-03	90
102	Willie Dryve	12.5000	South	Downtown	*NULL*	PXK3D7T	Chevy	Tahoe	L	12.5000	USB Port, Navigation	2020-04-05	80
103	Sal Debote	10.0000	North	Downtown	East	445GH2	Nissan	Leaf	S	7.5000	USB Port, XM Radio	2020-04-12	90
103	Sal Debote	10.0000	North	Downtown	East	59DLLK	Chevy	Trax	S	7.5000	USB, Bluetooth	2020-04-02	85
103	Sal Deboat	10.0000	North	Downtown	East	667GM8	Nissan	Altima	M	10.0000	USB Port,Blueooth,Naviation	2020-04-11	97
104	Carol Ling	12.5000	South	*NULL*	West	663ETMP	Chevy	Surburban	L	12.5000	XM Radio	2020-04-12	92
104	Carol Ling	12.5000	South	*NULL*	West	667GM8	Nissan	Altima	M	10.0000	USB Port,Blueooth,Naviation	2020-04-09	94
104	Carol Ling	12.5000	South	*NULL*	West	PPF673	Cadillac	Escalade	M	10.0000	USB Port,Navigation,XM Radio…	2020-04-04	83
105	Ida Knowe	5.0000	*NULL*	*NULL*	Downtown	445GH2	Nissan	Leaf	S	7.5000	USB Port, XM Radio	2020-04-17	99

Insert Anomaly

How do we add Kent Belevit? Not until he takes a test apparently! Then to insert the test is an SQL UPDATE?



driver_id	driver_name	driver_fee	region1	region2	region3	licplate	make	model	car_size	car_fee	car_features	test_date	test_score
101	Bill Melator	7.5000	West	North	Downtown	445GH2	Nissan	Leaf	S	7.5000	USB Port, XM Radio	2020-04-03	90
101	Bill Melator	7.5000	West	North	Downtown	59DLLK	Chevy	Trax	S	7.5000	USB, Bluetooth	2020-04-01	78
101	Bill Melator	7.5000	West	North	Downtown	PPF673	Cadillac	Escalade	M	10.0000	USB Port,Navigation,XM Radio…	2020-04-05	88
101	Bill Melator	7.5000	West	North	Downtown	PXK3D7T	Chevy	Tahoe	L	12.5000	USB Port, Navigation	2020-04-06	92
102	Willie Dryve	12.5000	South	Downtown	*NULL*	663ETMP	Chevy	Surburban	L	12.5000	XM Radio	2020-04-03	90
102	Willie Dryve	12.5000	South	Downtown	*NULL*	PXK3D7T	Chevy	Tahoe	L	12.5000	USB Port, Navigation		80
103	Sal Debote	10.0000	North	Downtown	East	445GH2	Nisaan	Leaf	S	7.5000	USB Port, XM Radio		90
103	Sal Debote	10.0000	North	Downtown	East	59DLLK	Chevy	Trax	S	7.5000	USB, Bluetooth		85
103	Sal Deboat	10.0000	North	Downtown	East	667GM8	Nissan	Altima	M	10.0000	USB Port,Blueooth,Navi		97
104	Carol Ling	12.5000	South	*NULL*	West	663ETMP	Chevy	Surburban	L	12.5000	XM Radio	2020-04-12	92
104	Carol Ling	12.5000	South	*NULL*	West	667GM8	Nissan	Altima	M	10.0000	USB Port,Blueooth,Naviation	2020-04-09	94
104	Carol Ling	12.5000	South	*NULL*	West	PPF673	Cadillac	Escalade	M	10.0000	USB Port,Navigation,XM Radio…	2020-04-04	83
105	Ida Knowe	5.0000	*NULL*	*NULL*	Downtown	445GH2	Nissan	Leaf	S	7.5000	USB Port, XM Radio	2020-04-17	99

Nulls not allowed here!

Syracuse University
School of Information Studies

Data Anomalies

The End

Syracuse University
School of Information Studies

Summary of Data Anomalies

Back to the DRY Rule of Table Design

- One row per entity
  - The same thing should not appear more than once
  - For example: no table with the same customer in there three times
- The table should contain attributes directly related to it
  - Unrelated attributes in a separate table
  - Don’t hide relationships inside the table
  - For example: don’t include order data in the customer table
- Easier said than done!

What’s Next?

- We have:
  - Data redundancy
  - Data anomalies
  - Data inconsistencies
- Caused by poor design
- So, how do we fix it?
- Answer: normalize
  - Identify the data dependencies, then resolve them!

Syracuse University
School of Information Studies

Summary of Data Anomalies

The End

Syracuse University
School of Information Studies

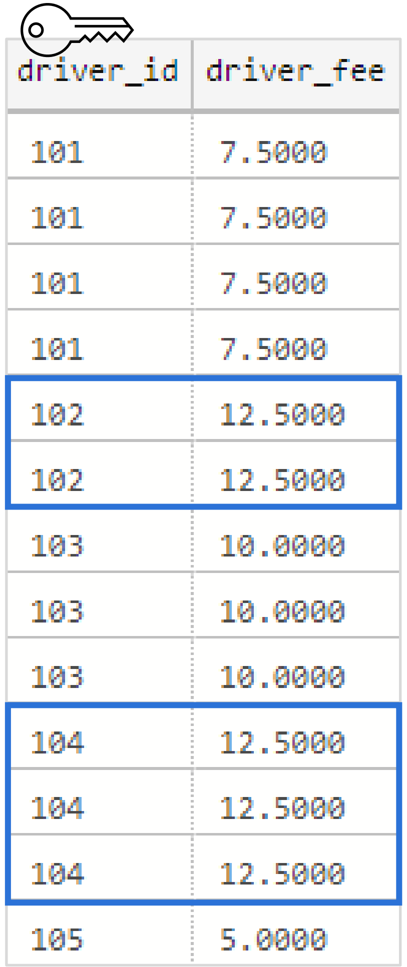
Data Dependencies

Data Dependencies

- Data dependencies define the data relationship between the attributes in two columns of a table
- Each column might:
  - Not be useable by the key
  - Act as a key
  - Be a key
  - Be part of a key

Key Dependency

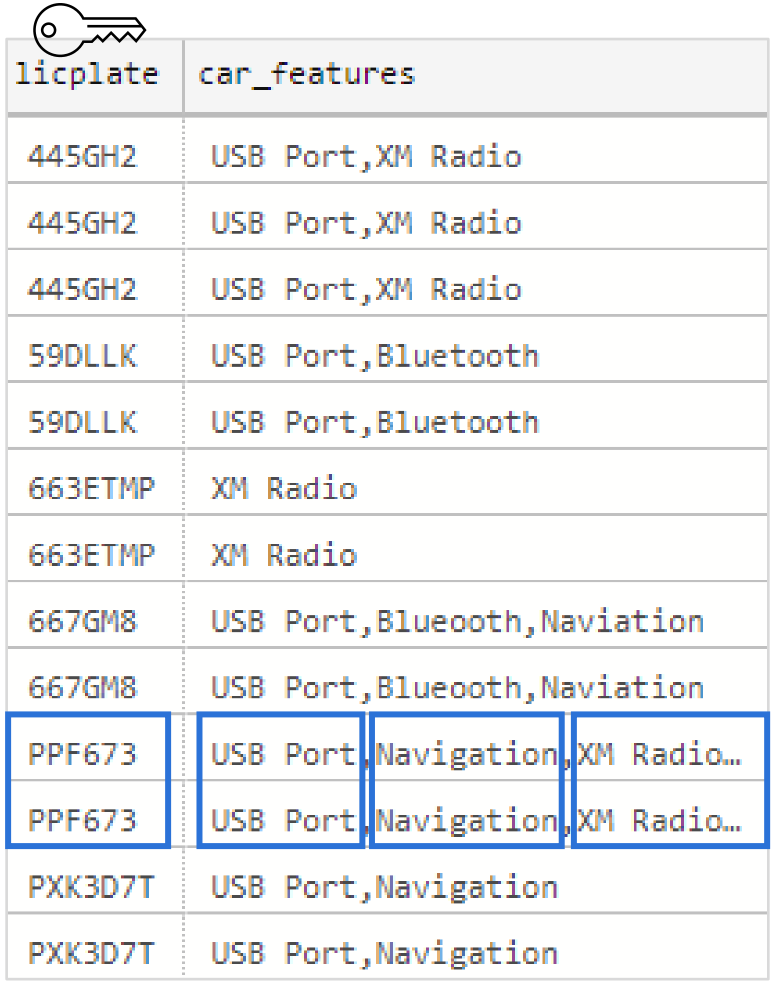
- Key dependency exists between columns A and B when column A is a key and can be used to lookup a single atomic value for column B.
- With driver_id as the key, we can look up a single atomic value for driver_fee.
- While the key does repeat, the same value is looked up for each key.



driver_id	driver_fee
101	7.5000
101	7.5000
101	7.5000
101	7.5000
102	12.5000
102	12.5000
103	10.0000
103	10.0000
103	10.0000
104	12.5000
104	12.5000
104	12.5000
105	5.0000

No Key Dependency

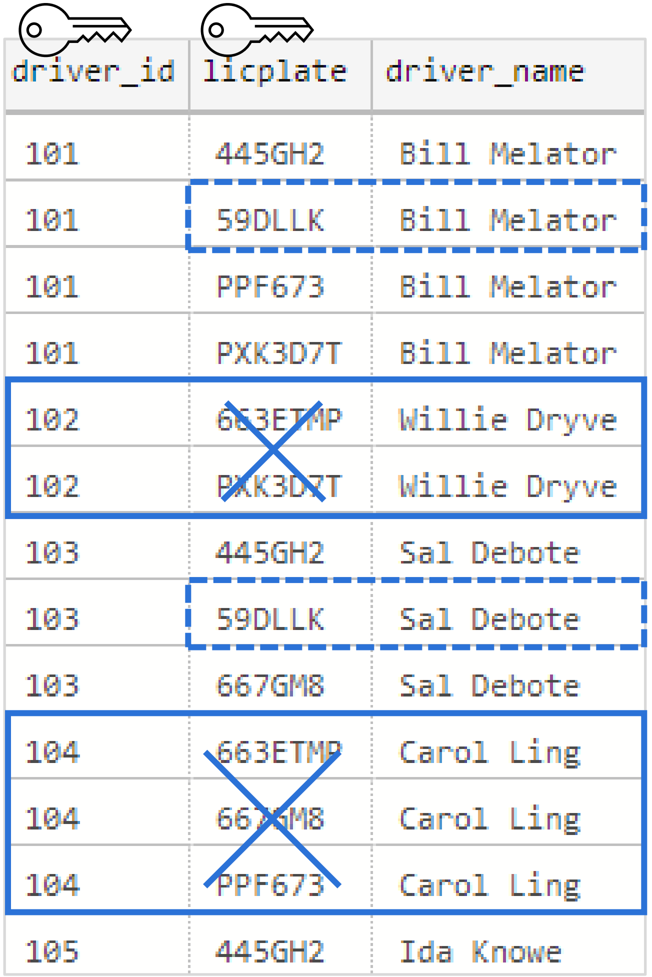
- No key dependency exists between columns A and B when column A is a key and cannot be used to lookup a single atomic value from column B.
- In this example, the key licplate cannot be used to look up a single feature (not atomic, but multivalued).
- While the same licplate looks up the same features, you cannot select a single atomic feature like XM Radio.



licplate	car_features
445GH2	USB Port,XM Radio
445GH2	USB Port,XM Radio
445GH2	USB Port,XM Radio
59DLLK	USB Port,Bluetooth
59DLLK	USB Port,Bluetooth
663ETMP	XM Radio
663ETMP	XM Radio
667GM8	USB Port,Blueooth,Naviation
667GM8	USB Port,Blueooth,Naviation
PPF673	USB Port,Navigation,XM Radio…
PPF673	USB Port,Navigation,XM Radio…
PXK3D7T	USB Port,Navigation
PXK3D7T	USB Port,Navigation

Partial Key Dependency

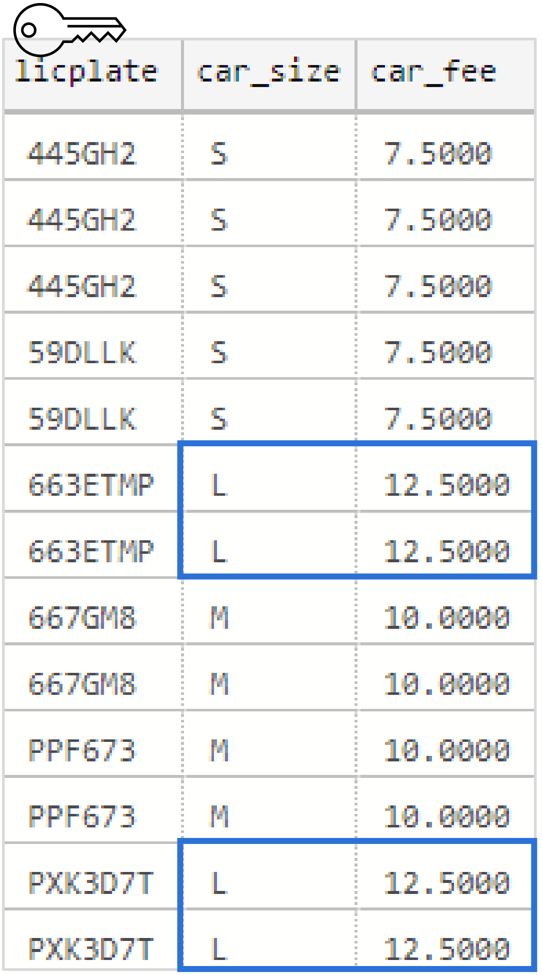
- A partial key dependency exists between columns A and B when columns A and C are a key and only column A is only necessary to look up a single atomic value for column B.
- Columns driver_id and licplate are the key, but only the driver_id is required to look up single atomic values of driver_name.



driver_id	licplate	driver_name
101	445GH2	Bill Melator
101	59DLLK	Bill Melator
101	PPF673	Bill Melator
101	PXK3D7T	Bill Melator
102	~~663ETMP~~	Willie Dryve
102	~~PXK3D7T~~	Willie Dryve
103	445GH2	Sal Debote
103	59DLLK	Sal Debote
103	667GM8	Sal Debote
104	~~663ETMP~~	Carol Ling
104	~~667GM8~~	Carol Ling
104	~~PPF673~~	Carol Ling
105	445GH2	Ida Knowe

Transitive Dependency

- A transitive dependency exists between columns A and B when column A is not a key but can be used to look up a single atomic value in column B.
- In this example, licplate is the key, but car_size can be used to look up a single atomic value for car_fee.
- Column car_size is acting as a key on behalf of licplate, hence the transitive dependency.



licplate	car_size	car_fee
445GH2	S	7.5000
445GH2	S	7.5000
445GH2	S	7.5000
59DLLK	S	7.5000
59DLLK	S	7.5000
663ETMP	L	12.5000
663ETMP	L	12.5000
667GM8	M	10.0000
667GM8	M	10.0000
PPF673	M	10.0000
PPF673	M	10.0000
PXK3D7T	L	12.5000
PXK3D7T	L	12.5000

Syracuse University
School of Information Studies

Data Dependencies

The End

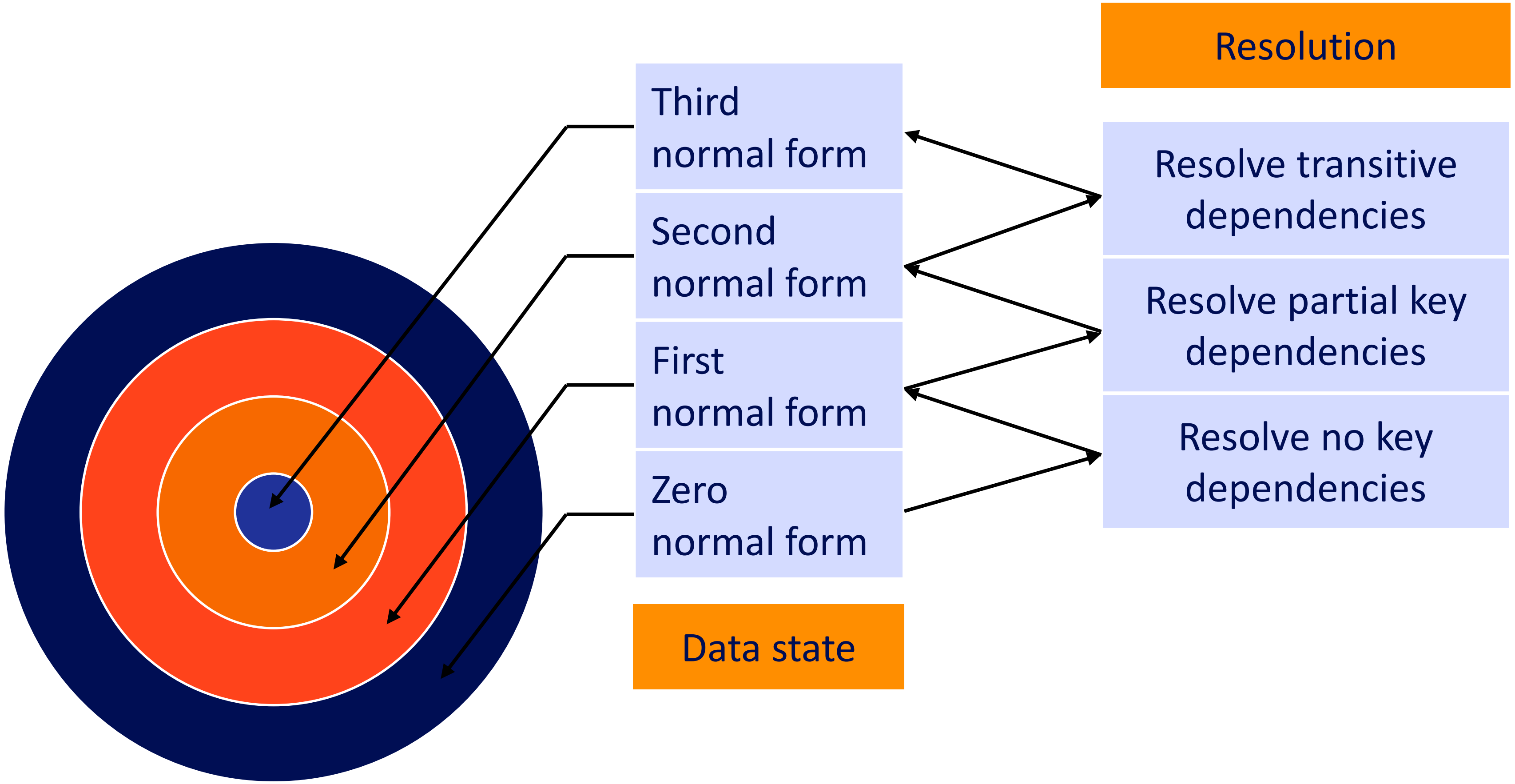
Syracuse University
School of Information Studies

Normal Forms

Normal Form

- A normal form represents the state of the data within the table with respect to data dependencies.
- Zero normal form (0NF) data have at least one column with no key dependency.
- First normal form (1NF) has zero columns with no key dependency.
- Second normal form (2NF) has zero columns with no key dependency and zero columns with partial key dependency.
- Third normal form (3NF) has zero columns with no key dependency, zero columns with partial key dependency, and zero columns with transitive dependency.

Progression of Normal Forms



Resolution

Third normal form

Second normal form

First normal form

Zero normal form

Data state

Resolve transitive dependencies

Resolve partial key dependencies

Resolve no key dependencies

Syracuse University
School of Information Studies

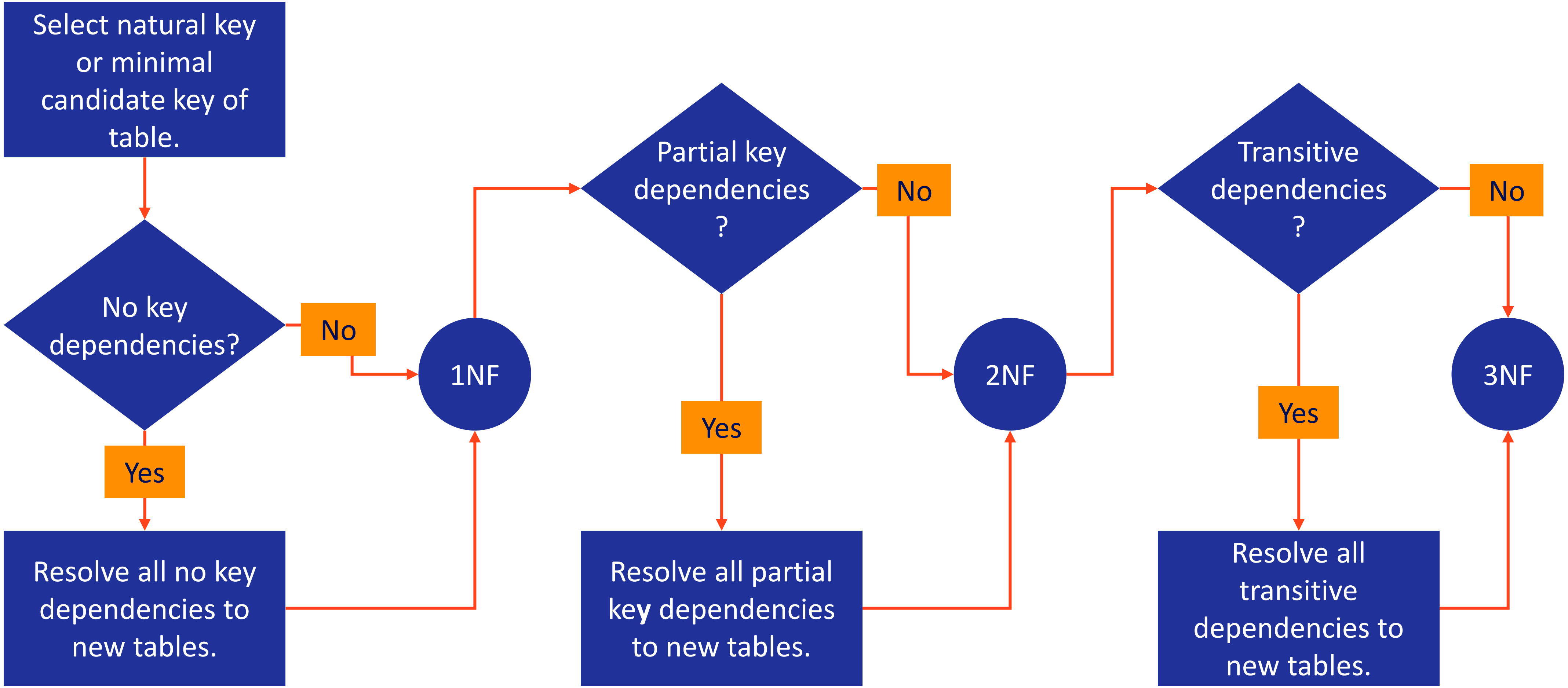
Normal Forms

The End

Syracuse University
School of Information Studies

The Data Normalization Process

Data Normalization Process



Select natural key or minimal candidate key of table.
No key dependencies?
No
Yes
Resolve all no key dependencies to new tables.
1NF
Partial key dependencies ?
No
Yes
Resolve all partial key dependencies to new tables.
2NF
Transitive dependencies ?
No
Yes
Resolve all transitive dependencies to new tables.
3NF

Syracuse University
School of Information Studies

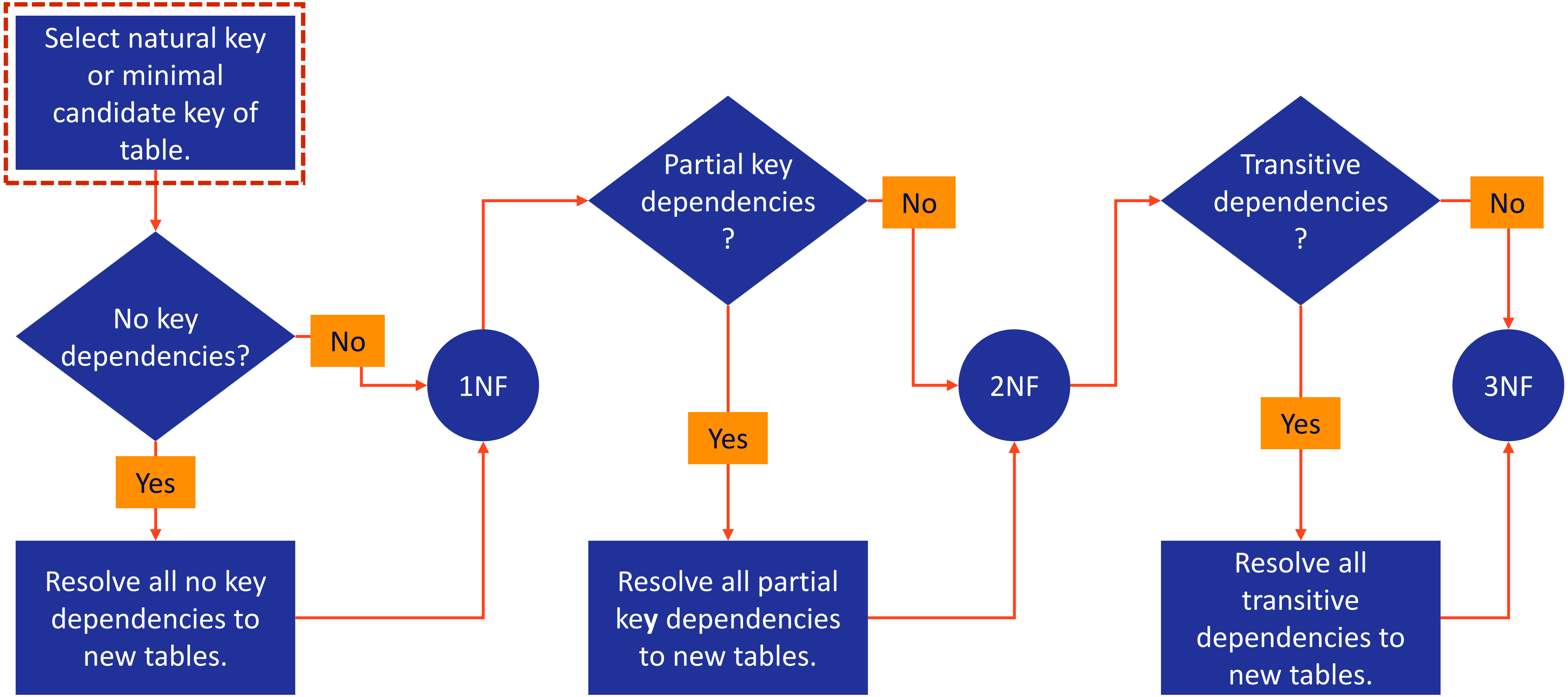
The Data Normalization Process

The End

Syracuse University
School of Information Studies

Key Selection

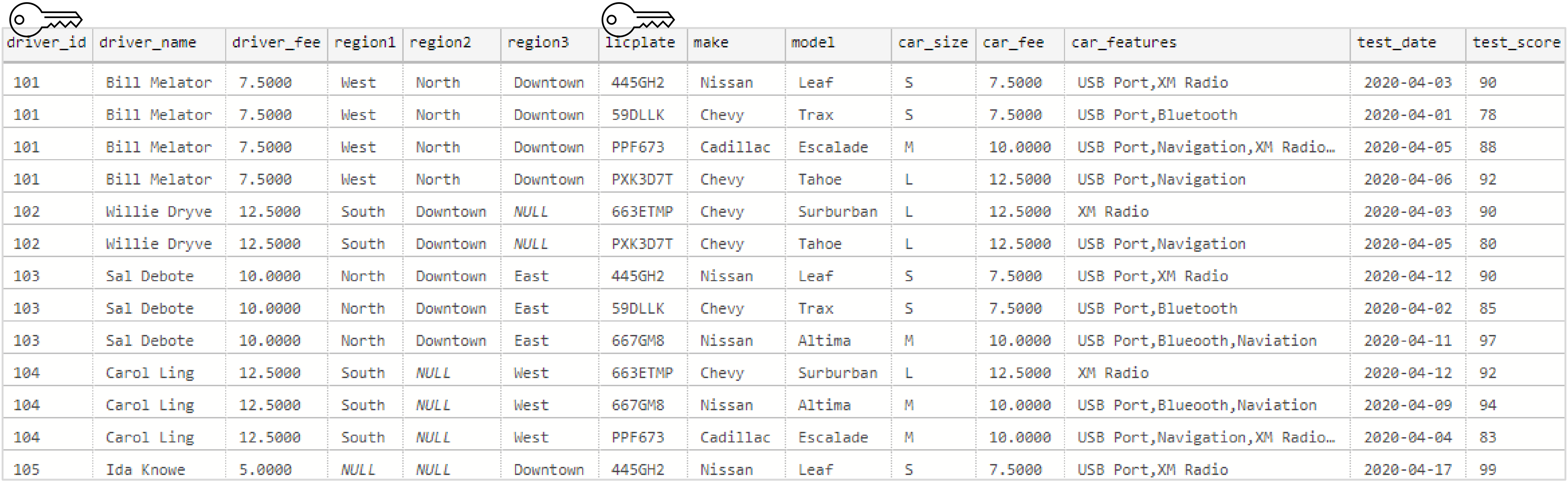
Key Selection


Select natural key or minimal candidate key of table.
No key dependencies?
No
Yes
Resolve all no key dependencies to new tables.
1NF
Partial key dependencies ?
No
Yes
Resolve all partial key dependencies to new tables.
2NF
Transitive dependencies ?
No
Yes
Resolve all transitive dependencies to new tables.
3NF

Key Selection (cont.)

- For each table to be normalized, a key must be selected from the table data.
- This key cannot be a surrogate key. Since these keys are meaningless, they have no data dependencies.
- A natural or business key should be selected.
- If there is no natural key, the minimal candidate key should be selected.

Cuse Rides: Key Selection



🔑 driver_id	driver_name	driver_fee	region1	region2	region3	🔑 licplate	make	model	car_size	car_fee	car_features	test_date	test_score
101	Bill Melator	7.5000	West	North	Downtown	445GH2	Nissan	Leaf	S	7.5000	USB Port,XM Radio	2020-04-03	90
101	Bill Melator	7.5000	West	North	Downtown	59DLLK	Chevy	Trax	S	7.5000	USB Port,Bluetooth	2020-04-01	78
101	Bill Melator	7.5000	West	North	Downtown	PPF673	Cadillac	Escalade	M	10.0000	USB Port,Navigation,XM Radio…	2020-04-05	88
101	Bill Melator	7.5000	West	North	Downtown	PXK3D7T	Chevy	Tahoe	L	12.5000	USB Port,Navigation	2020-04-06	92
102	Willie Dryve	12.5000	South	Downtown	*NULL*	663ETMP	Chevy	Surburban	L	12.5000	XM Radio	2020-04-03	90
102	Willie Dryve	12.5000	South	Downtown	*NULL*	PXK3D7T	Chevy	Tahoe	L	12.5000	USB Port,Navigation	2020-04-05	80
103	Sal Debote	10.0000	North	Downtown	East	445GH2	Nissan	Leaf	S	7.5000	USB Port,XM Radio	2020-04-12	90
103	Sal Debote	10.0000	North	Downtown	East	59DLLK	Chevy	Trax	S	7.5000	USB Port,Bluetooth	2020-04-02	85
103	Sal Debote	10.0000	North	Downtown	East	667GM8	Nissan	Altima	M	10.0000	USB Port,Blueooth,Naviation	2020-04-11	97
104	Carol Ling	12.5000	South	*NULL*	West	663ETMP	Chevy	Surburban	L	12.5000	XM Radio	2020-04-12	92
104	Carol Ling	12.5000	South	*NULL*	West	667GM8	Nissan	Altima	M	10.0000	USB Port,Blueooth,Naviation	2020-04-09	94
104	Carol Ling	12.5000	South	*NULL*	West	PPF673	Cadillac	Escalade	M	10.0000	USB Port,Navigation,XM Radio…	2020-04-04	83
105	Ida Knowe	5.0000	*NULL*	*NULL*	Downtown	445GH2	Nissan	Leaf	S	7.5000	USB Port,XM Radio	2020-04-17	99

One row represents a driver passing a vehicle exam. It does not make sense for the driver id and license plate to repeat.

Syracuse University
School of Information Studies

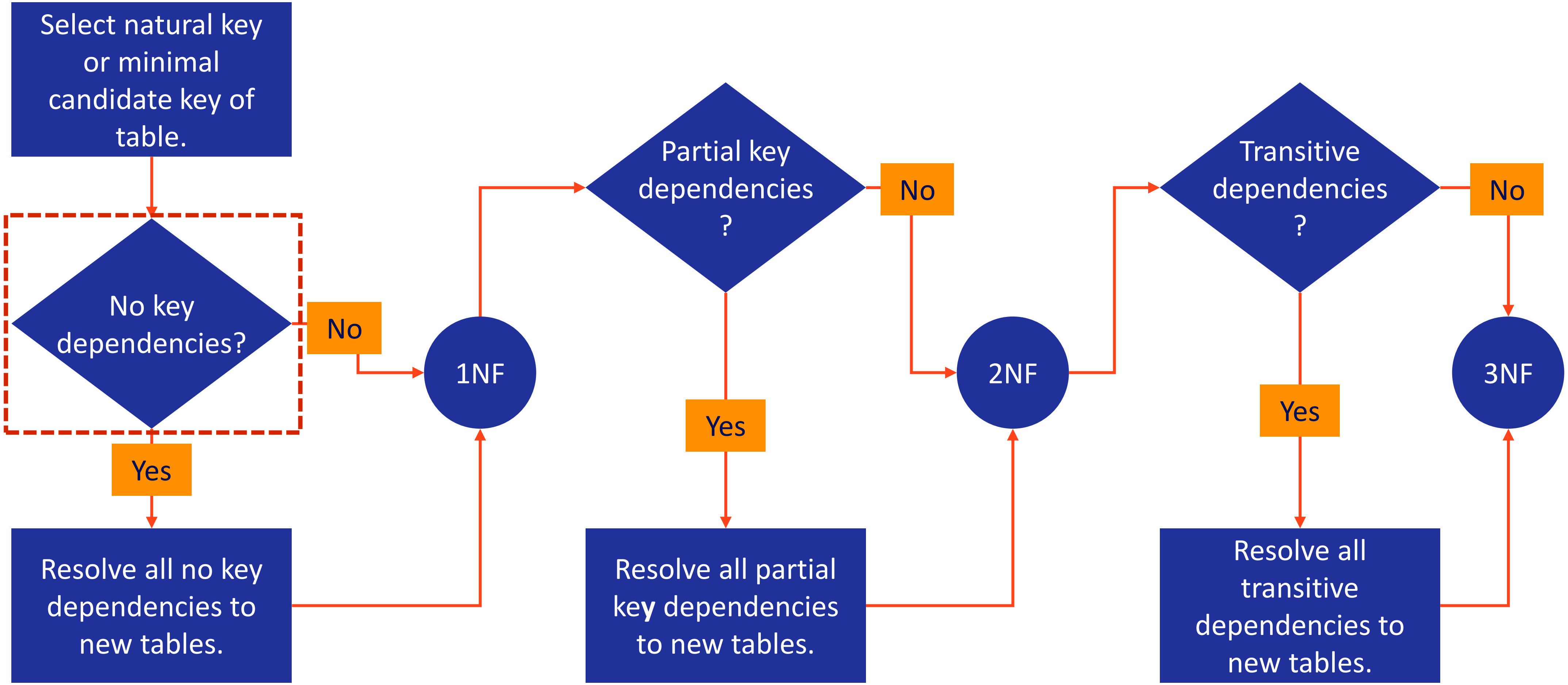
Key Selection

The End

Syracuse University
School of Information Studies

Identifying No Key Dependencies

Identifying No Key Dependencies



Select natural key or minimal candidate key of table.

No key dependencies?

No

Yes

Resolve all no key dependencies to new tables.

1NF

Partial key dependencies ?

No

Yes

Resolve all partial key dependencies to new tables.

2NF

Transitive dependencies ?

No

Yes

Resolve all transitive dependencies to new tables.

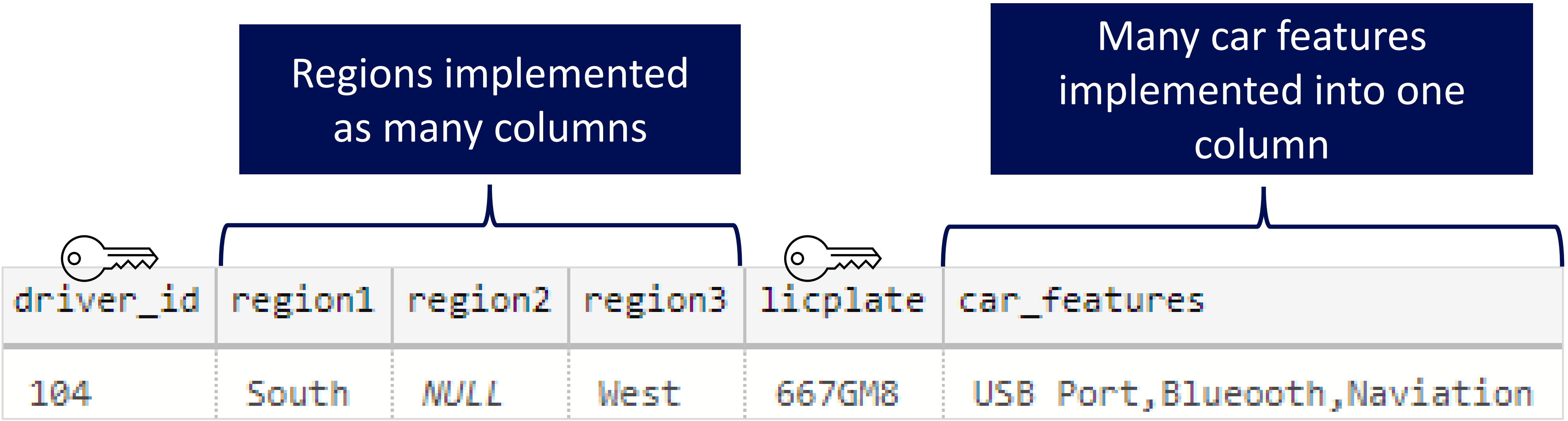
3NF

Identifying No Key Dependencies

- No key dependencies are multivalued attributes that were implemented in the relational table as multiple values in a single column or multiple columns with similar names.
- In both cases, the data are not query-able in the current format because you must inspect inside the field or know which of many fields to choose.
- From the data modeling perspective, consider no key dependencies as [M] attributes that were never resolved to an M-M relationship.

Cuse Rides: No Key Dependencies

- The Cuse Rides fleet table has two no key dependencies.
- A driver drives in many regions, such as the west side or downtown.
- A car is equipped with many convenience features, such as USB ports or Bluetooth.



Regions implemented as many columns

Many car features implemented into one column

driver_id	region1	region2	region3	licplate	car_features
104	South	*NULL*	West	667GM8	USB Port,Blueooth,Naviation

Syracuse University
School of Information Studies

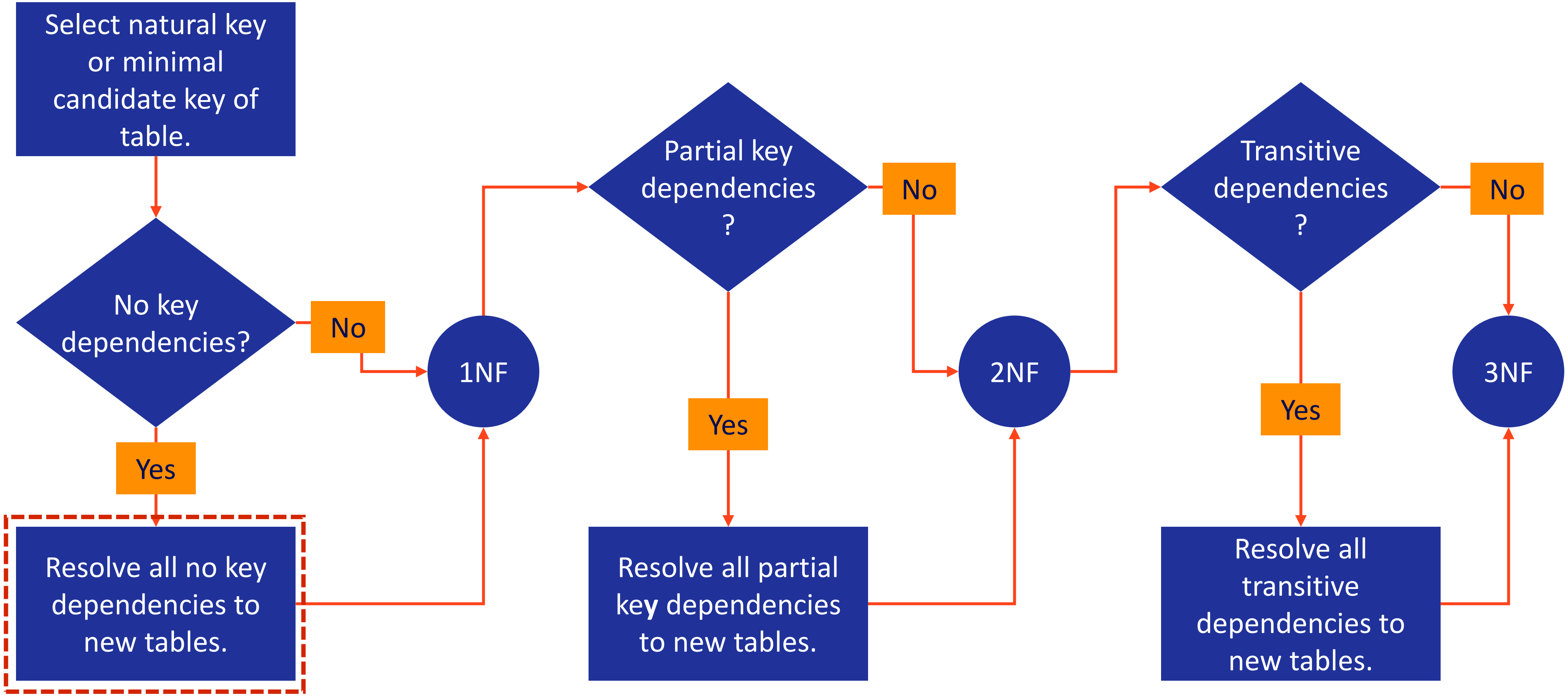
Identifying No Key Dependencies

The End

Syracuse University
School of Information Studies

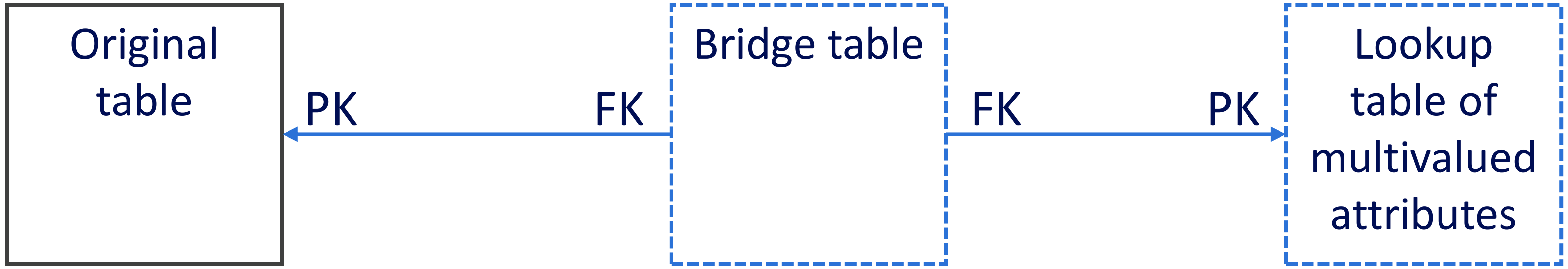
Resolving No Key Dependencies

Resolve No Key Dependencies

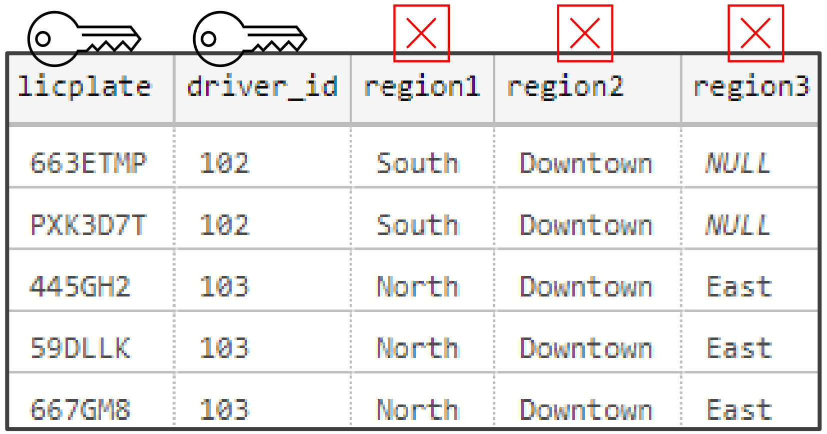

Select natural key or minimal candidate key of table.
No key dependencies?
No
Yes
Resolve all no key dependencies to new tables.
1NF
Partial key dependencies ?
No
Yes
Resolve all partial key dependencies to new tables.
2NF
Transitive dependencies ?
No
Yes
Resolve all transitive dependencies to new tables.
3NF

Resolving No Key Dependencies

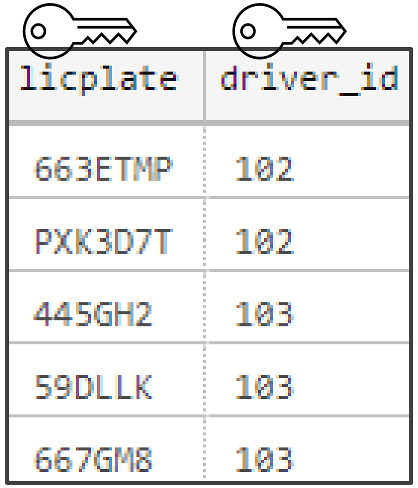
- These are many-to-many relationships, hidden inside a single column or group of columns.
- So, we need to add a lookup table and a bridge table.
- Lookup table holds the multivalues; bridge table sustains the original relationship to the key.
- No key dependencies are moved to the bridge table, while unique values go in the lookup table.



Cuse Rides: Resolve Regions

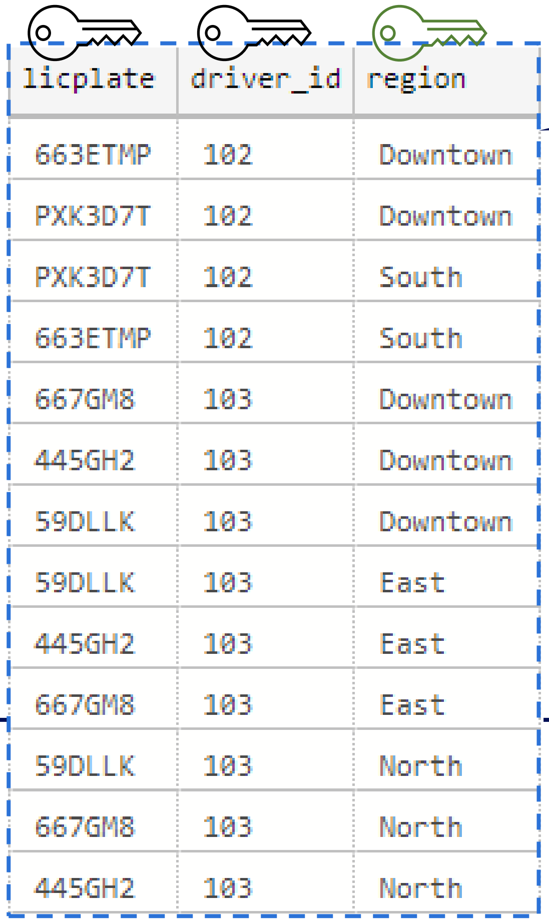


licplate	driver_id	region1	region2	region3
663ETMP	102	South	Downtown	*NULL*
PXK3D7T	102	South	Downtown	*NULL*
445GH2	103	North	Downtown	East
59DLLK	103	North	Downtown	East
667GM8	103	North	Downtown	East



licplate	driver_id
663ETMP	102
PXK3D7T	102
445GH2	103
59DLLK	103
667GM8	103

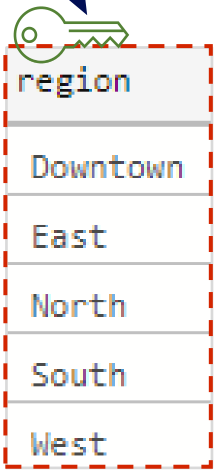
Dependencies removed



licplate	driver_id	region
663ETMP	102	Downtown
PXK3D7T	102	Downtown
PXK3D7T	102	South
663ETMP	102	South
667GM8	103	Downtown
445GH2	103	Downtown
59DLLK	103	Downtown
59DLLK	103	East
445GH2	103	East
667GM8	103	East
59DLLK	103	North
667GM8	103	North
445GH2	103	North

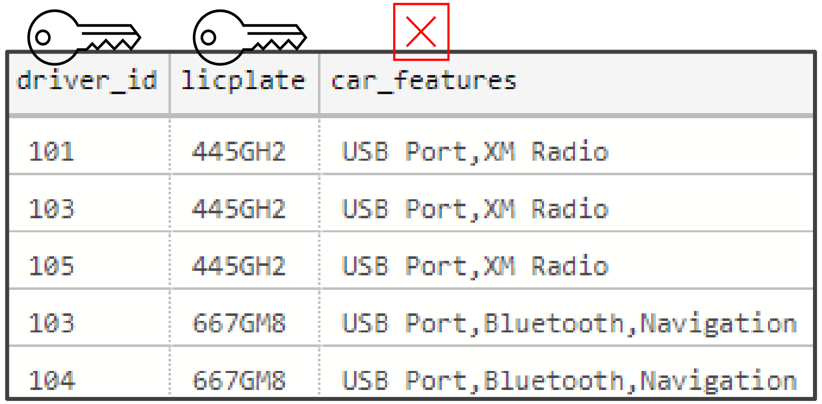
Bridge table

Lookup table

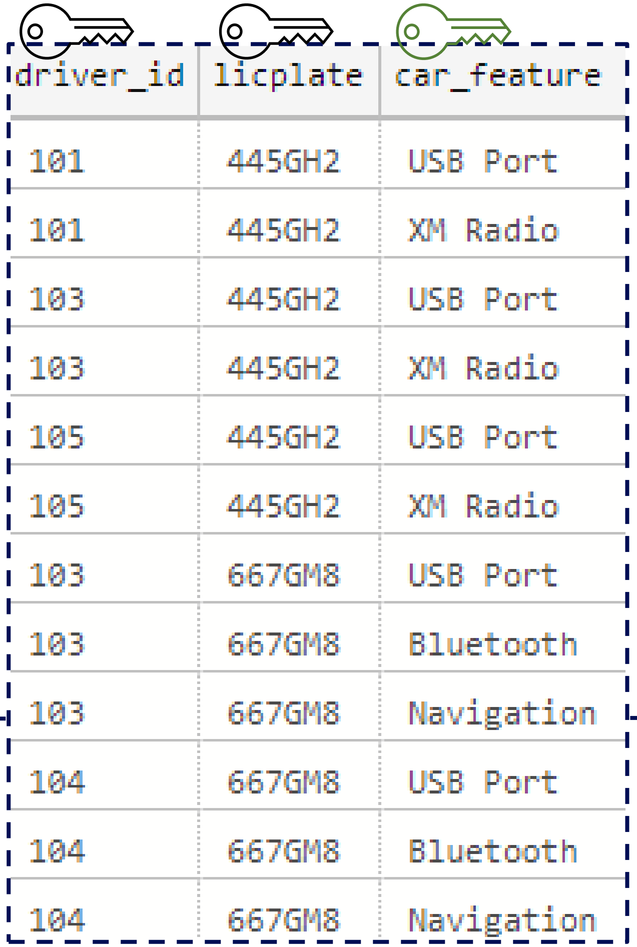


region
Downtown
East
North
South
West

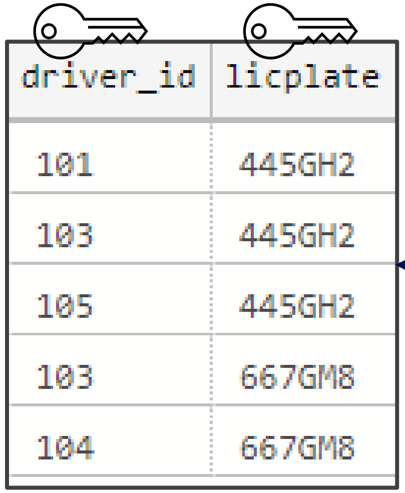
Cuse Rides: Resolve Car Features



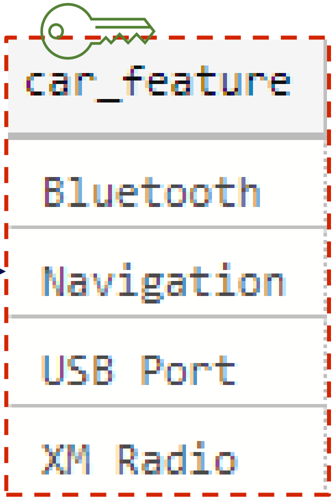
driver_id	licplate	car_features
101	445GH2	USB Port,XM Radio
103	445GH2	USB Port,XM Radio
105	445GH2	USB Port,XM Radio
103	667GM8	USB Port,Bluetooth,Navigation
104	667GM8	USB Port,Bluetooth,Navigation



driver_id	licplate	car_feature
101	445GH2	USB Port
101	445GH2	XM Radio
103	445GH2	USB Port
103	445GH2	XM Radio
105	445GH2	USB Port
105	445GH2	XM Radio
103	667GM8	USB Port
103	667GM8	Bluetooth
103	667GM8	Navigation
104	667GM8	USB Port
104	667GM8	Bluetooth
104	667GM8	Navigation



driver_id	licplate
101	445GH2
103	445GH2
105	445GH2
103	667GM8
104	667GM8



car_feature
Bluetooth
Navigation
USB Port
XM Radio

Syracuse University
School of Information Studies

Resolving No Key Dependencies

The End

Syracuse University
School of Information Studies

Demo

Resolve No Key Dependencies

Demo: Resolve No Key Dependencies

- We will use the Azure Data Studio application.
- We will use the demo database, cr_fleet table.
- Create the tables.
  - cr_regions (lookup table)
  - cr_fleet_regions (bridge table)
  - cr_features (lookup table)
  - cr_fleet_features (bridge table)
  - Don’t do the FKs until we are done normalizing
- Migrate data.
  - UNPIVOT for multiple columns region1, region2, region3
  - STRING_SPLIT for single-column car_features
  - INSERT INTO SELECT, SELECT INTO

Syracuse University
School of Information Studies

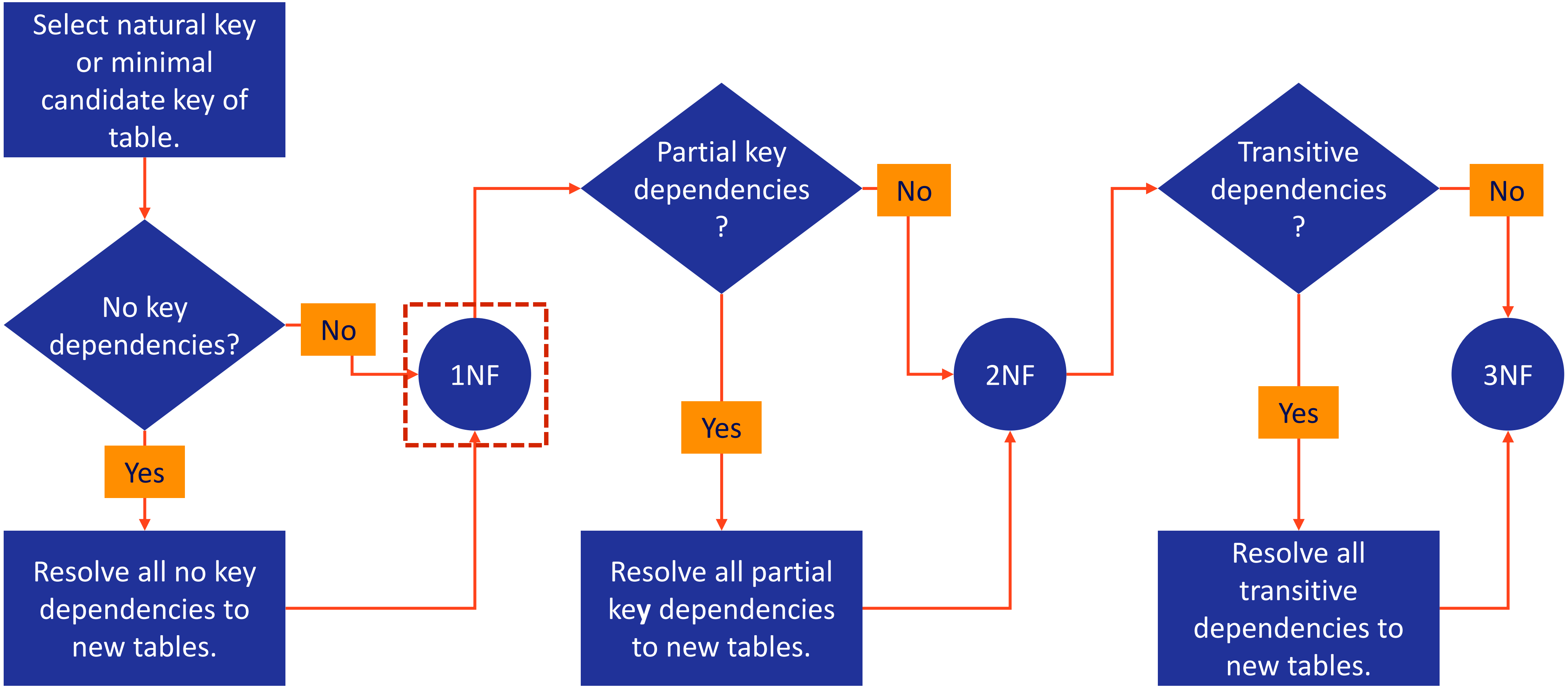
Demo: Resolve No Key Dependencies

The End

Syracuse University
School of Information Studies

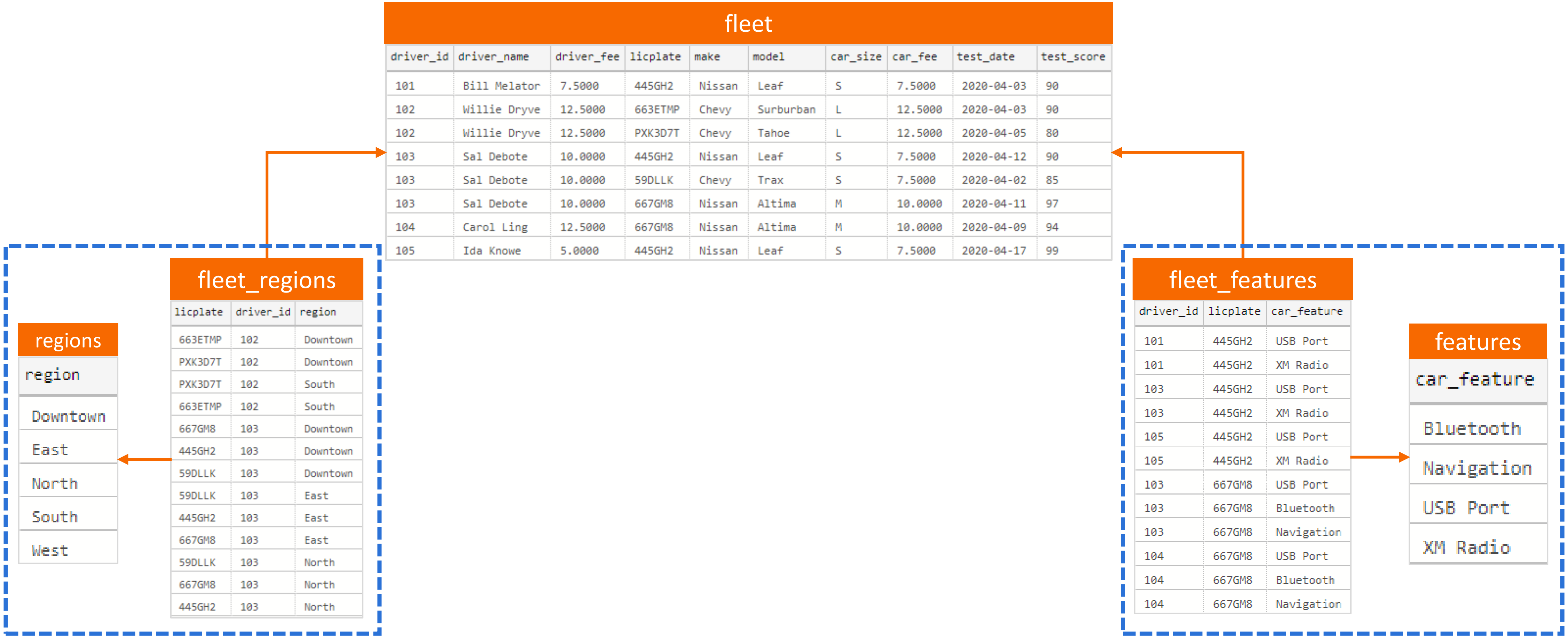
First Normal Form

First Normal Form



Select natural key or minimal candidate key of table.
No key dependencies?
No
Yes
Resolve all no key dependencies to new tables.
1NF
Partial key dependencies ?
No
Yes
Resolve all partial key dependencies to new tables.
2NF
Transitive dependencies ?
No
Yes
Resolve all transitive dependencies to new tables.
3NF

Cuse Rides: Resolved No Key Dependencies, Now in 1NF



fleet

driver_id	driver_name	driver_fee	licplate	make	model	car_size	car_fee	test_date	test_score
101	Bill Melator	7.5000	445GH2	Nissan	Leaf	S	7.5000	2020-04-03	90
102	Willie Dryve	12.5000	663ETMP	Chevy	Surburban	L	12.5000	2020-04-03	90
102	Willie Dryve	12.5000	PXK3D7T	Chevy	Tahoe	L	12.5000	2020-04-05	80
103	Sal Debote	10.0000	445GH2	Nissan	Leaf	S	7.5000	2020-04-12	90
103	Sal Debote	10.0000	59DLLK	Chevy	Trax	S	7.5000	2020-04-02	85
103	Sal Debote	10.0000	667GM8	Nissan	Altima	M	10.0000	2020-04-11	97
104	Carol Ling	12.5000	667GM8	Nissan	Altima	M	10.0000	2020-04-09	94
105	Ida Knowe	5.0000	445GH2	Nissan	Leaf	S	7.5000	2020-04-17	99

regions

region
Downtown
East
North
South
West

fleet_regions

licplate	driver_id	region
663ETMP	102	Downtown
PXK3D7T	102	Downtown
PXK3D7T	102	South
663ETMP	102	South
667GM8	103	Downtown
445GH2	103	Downtown
59DLLK	103	Downtown
59DLLK	103	East
445GH2	103	East
667GM8	103	East
59DLLK	103	North
667GM8	103	North
445GH2	103	North

fleet_features

driver_id	licplate	car_feature
101	445GH2	USB Port
101	445GH2	XM Radio
103	445GH2	USB Port
103	445GH2	XM Radio
105	445GH2	USB Port
105	445GH2	XM Radio
103	667GM8	USB Port
103	667GM8	Bluetooth
103	667GM8	Navigation
104	667GM8	USB Port
104	667GM8	Bluetooth
104	667GM8	Navigation

features

car_feature
Bluetooth
Navigation
USB Port
XM Radio

Syracuse University
School of Information Studies

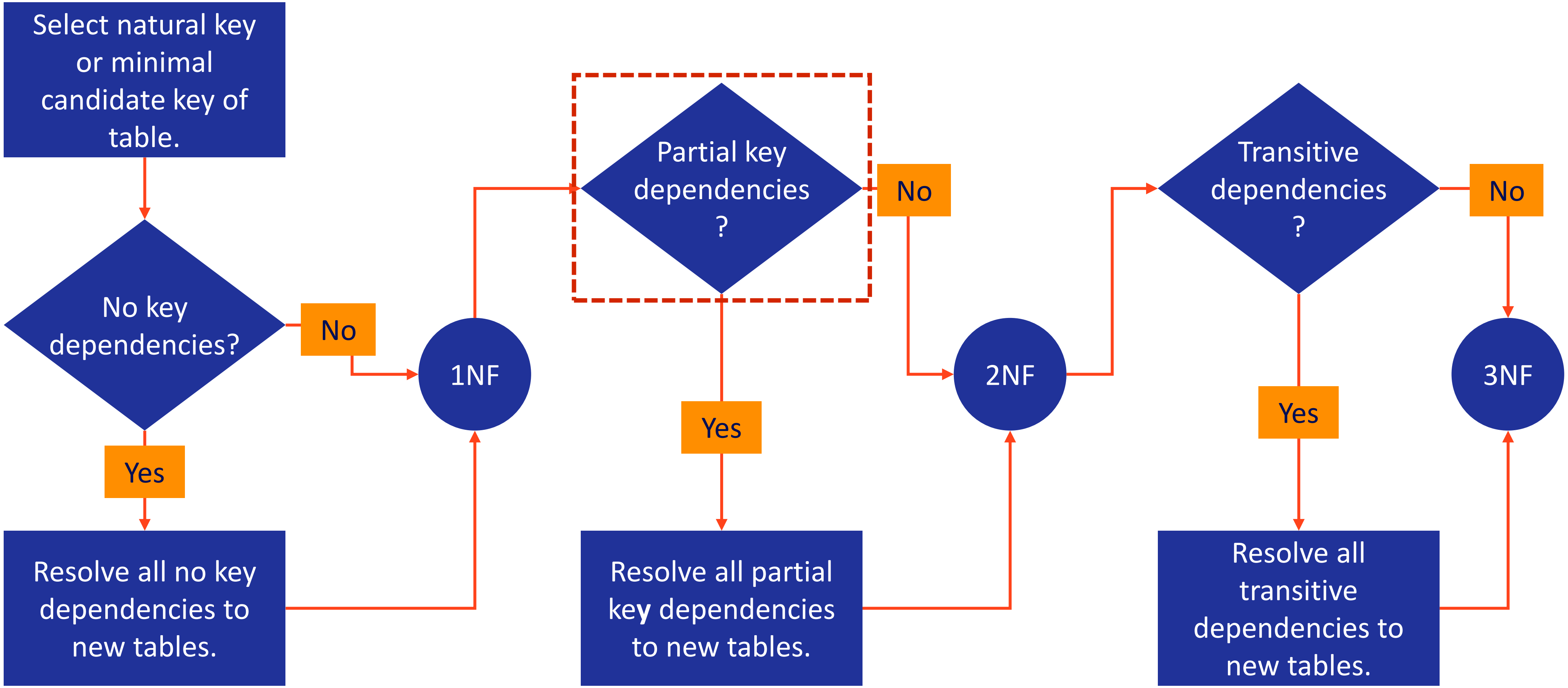
First Normal Form

The End

Syracuse University
School of Information Studies

Identifying Partial Key Dependencies

Identifying Partial Key Dependencies



Select natural key or minimal candidate key of table.

No key dependencies?

Yes

No

Resolve all no key dependencies to new tables.

1NF

Partial key dependencies ?

Yes

No

Resolve all partial key dependencies to new tables.

2NF

Transitive dependencies ?

Yes

No

Resolve all transitive dependencies to new tables.

3NF

Identifying Partial Key Dependencies (cont.)

- Partial key dependencies occur when there is a composite key, and only a part of the key is necessary to look up a single atomic value.
- These are usually many-to-many relationships embedded within a single table.
- If the table has a single column key, it cannot have a partial dependency.
- If the table does not have any no key columns, it cannot have a partial dependency.

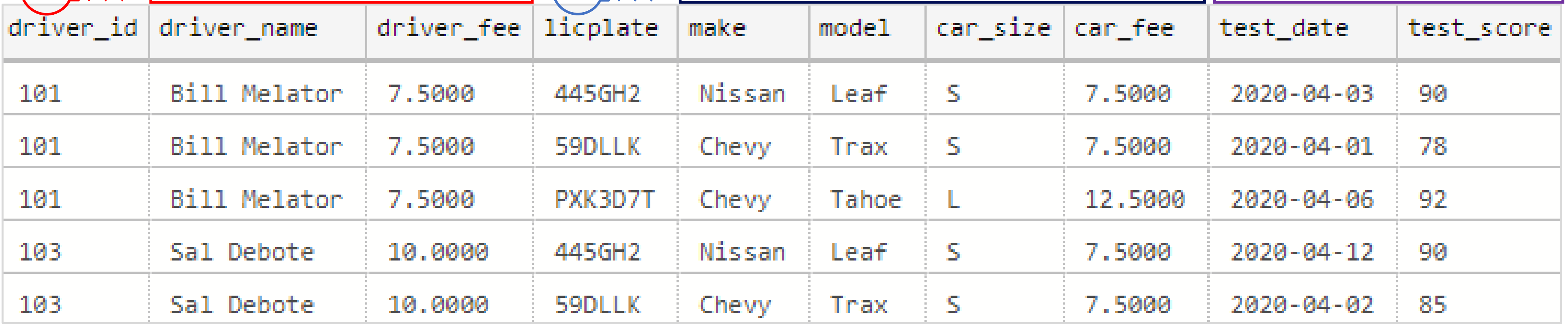
Cuse Rides: Partial Key Dependencies

- The Cuse Rides fleet table has two partial key dependencies.
- Only the driver_id is necessary to look up driver_name and driver_fee.
- Only the licplate is needed to look up make, model, car_size, and car_fee.

Only needs driver_id

Only needs the licplate

Needs both!



driver_id	driver_name	driver_fee	licplate	make	model	car_size	car_fee	test_date	test_score
101	Bill Melator	7.5000	445GH2	Nissan	Leaf	S	7.5000	2020-04-03	90
101	Bill Melator	7.5000	59DLLK	Chevy	Trax	S	7.5000	2020-04-01	78
101	Bill Melator	7.5000	PXK3D7T	Chevy	Tahoe	L	12.5000	2020-04-06	92
103	Sal Debote	10.0000	445GH2	Nissan	Leaf	S	7.5000	2020-04-12	90
103	Sal Debote	10.0000	59DLLK	Chevy	Trax	S	7.5000	2020-04-02	85

Syracuse University
School of Information Studies

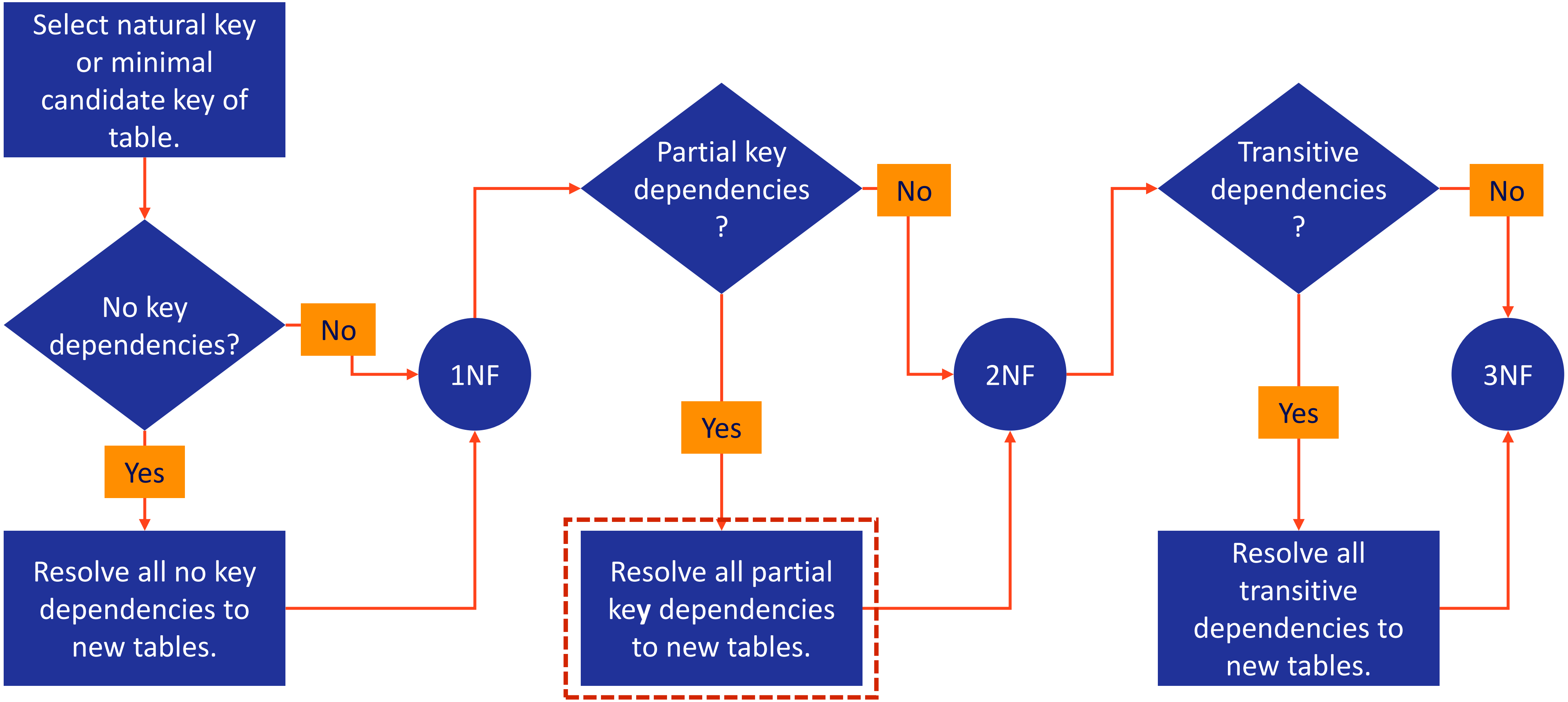
Identifying Partial Key Dependencies

The End

Syracuse University
School of Information Studies

Resolving Partial Key Dependencies

Resolving Partial Key Dependencies



Select natural key or minimal candidate key of table.

No key dependencies?

No

Yes

Resolve all no key dependencies to new tables.

1NF

Partial key dependencies ?

No

Yes

Resolve all partial key dependencies to new tables.

2NF

Transitive dependencies ?

No

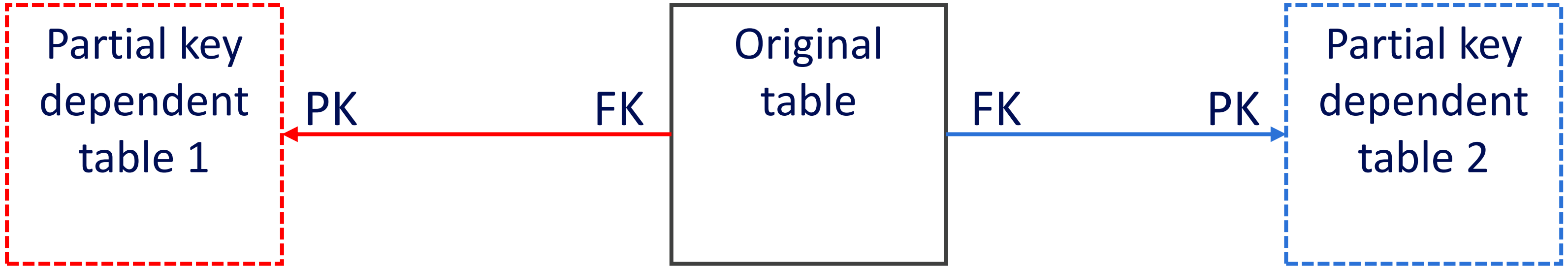
Yes

Resolve all transitive dependencies to new tables.

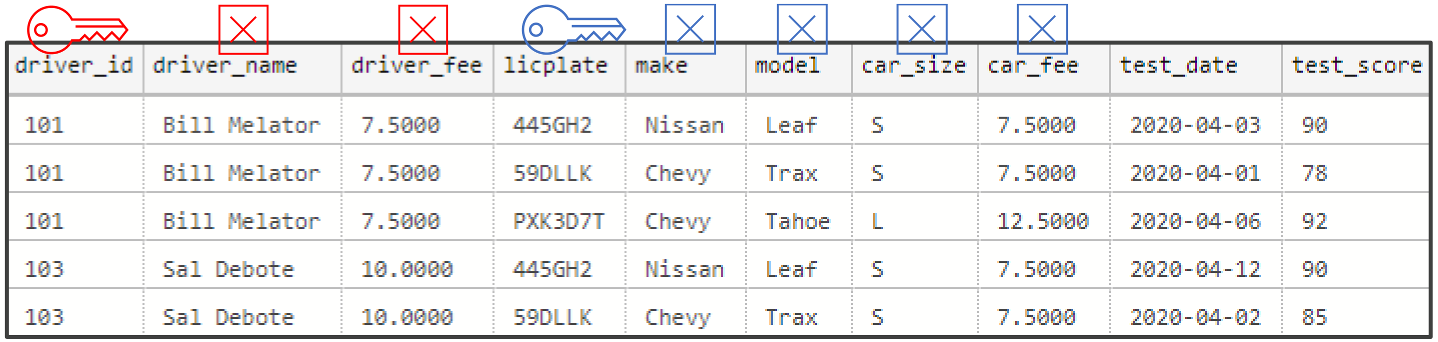
3NF

Resolving Partial Key Dependencies (cont.)

- Partially dependent no key values are removed from the original table and added to a new table.
- The primary key of the new table is the part of the key in which the columns are dependent.
- The original table is now an associative entity, with each part of the primary key now a foreign key to the new tables.

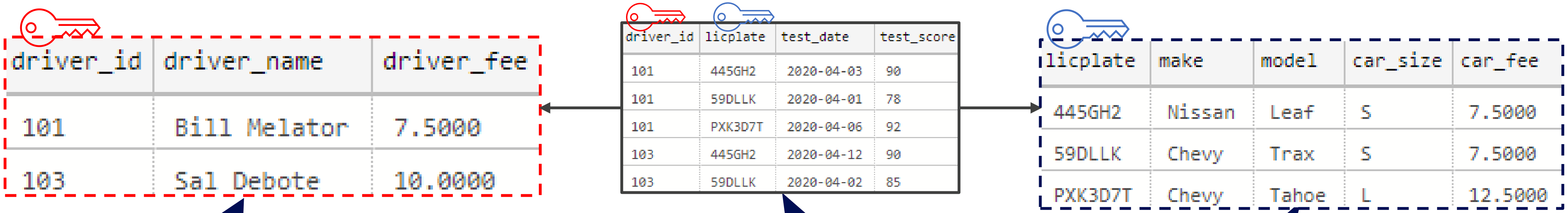


Cuse Rides: Resolve Partial Dependencies



driver_id	driver_name	driver_fee	licplate	make	model	car_size	car_fee	test_date	test_score
101	Bill Melator	7.5000	445GH2	Nissan	Leaf	S	7.5000	2020-04-03	90
101	Bill Melator	7.5000	59DLLK	Chevy	Trax	S	7.5000	2020-04-01	78
101	Bill Melator	7.5000	PXK3D7T	Chevy	Tahoe	L	12.5000	2020-04-06	92
103	Sal Debote	10.0000	445GH2	Nissan	Leaf	S	7.5000	2020-04-12	90
103	Sal Debote	10.0000	59DLLK	Chevy	Trax	S	7.5000	2020-04-02	85

Fleet table



driver_id	driver_name	driver_fee
101	Bill Melator	7.5000
103	Sal Debote	10.0000

driver_id	licplate	test_date	test_score
101	445GH2	2020-04-03	90
101	59DLLK	2020-04-01	78
101	PXK3D7T	2020-04-06	92
103	445GH2	2020-04-12	90
103	59DLLK	2020-04-02	85

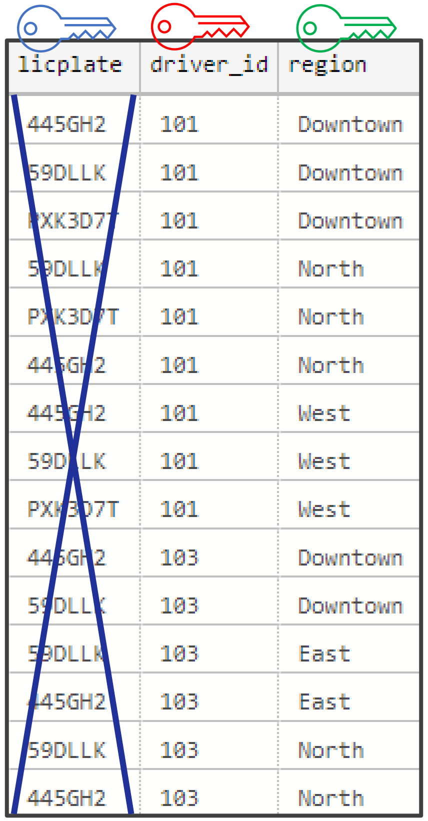
licplate	make	model	car_size	car_fee
445GH2	Nissan	Leaf	S	7.5000
59DLLK	Chevy	Trax	S	7.5000
PXK3D7T	Chevy	Tahoe	L	12.5000

Partial dependencies for driver_id

Dependencies removed

Partial dependencies for licplate

Cuse Rides: Resolve Partial Dependencies (cont.)

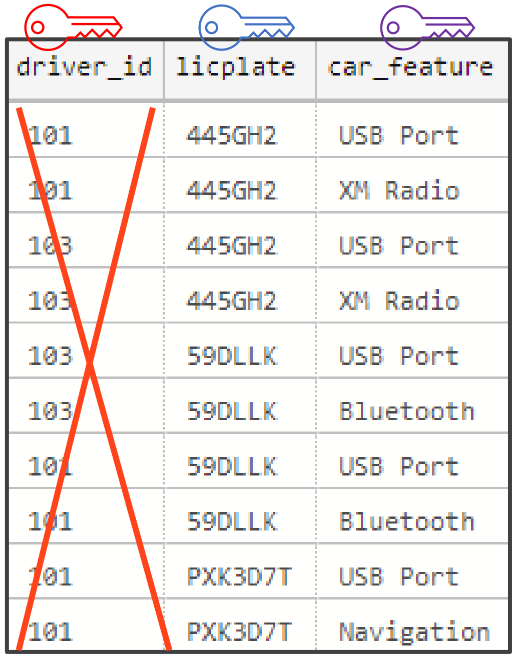


licplate	driver_id	region
445GH2	101	Downtown
59DLLK	101	Downtown
PXK3D7T	101	Downtown
59DLLK	101	North
PXK3D7T	101	North
445GH2	101	North
445GH2	101	West
59DLLK	101	West
PXK3D7T	101	West
445GH2	103	Downtown
59DLLK	103	Downtown
59DLLK	103	East
445GH2	103	East
59DLLK	103	North
445GH2	103	North

Now that we have a separate drivers table, in the fleet_regions table, we no longer need the licplate column. This table will now be joined to drivers.

Eliminate unnecessary columns from the key.

Now that we have a separate vehicles table, in the fleet_features table, we no longer require the **driver_id** column. This table will now be joined to vehicles.



driver_id	licplate	car_feature
101	445GH2	USB Port
101	445GH2	XM Radio
103	445GH2	USB Port
103	445GH2	XM Radio
103	59DLLK	USB Port
103	59DLLK	Bluetooth
101	59DLLK	USB Port
101	59DLLK	Bluetooth
101	PXK3D7T	USB Port
101	PXK3D7T	Navigation

Syracuse University
School of Information Studies

Resolving Partial Key Dependencies

The End

Syracuse University
School of Information Studies

Demo

Resolve Partial Key Dependencies

Demo: Resolve Partial Key Dependencies

- We will use the Azure Data Studio application.
- We will use the demo database, cr_fleet table.
- Create the tables.
  - cr_drivers (partial table)
  - cr_vehicles (partial table)
- Migrate data.
  - SELECT DISTINCT to reduce rows
  - INSERT INTO SELECT, SELECT INTO
- Alter cr_fleet_drivers, removing the redundant key.
- Alter cr_fleet_drivers, removing the redundant key.

Syracuse University
School of Information Studies

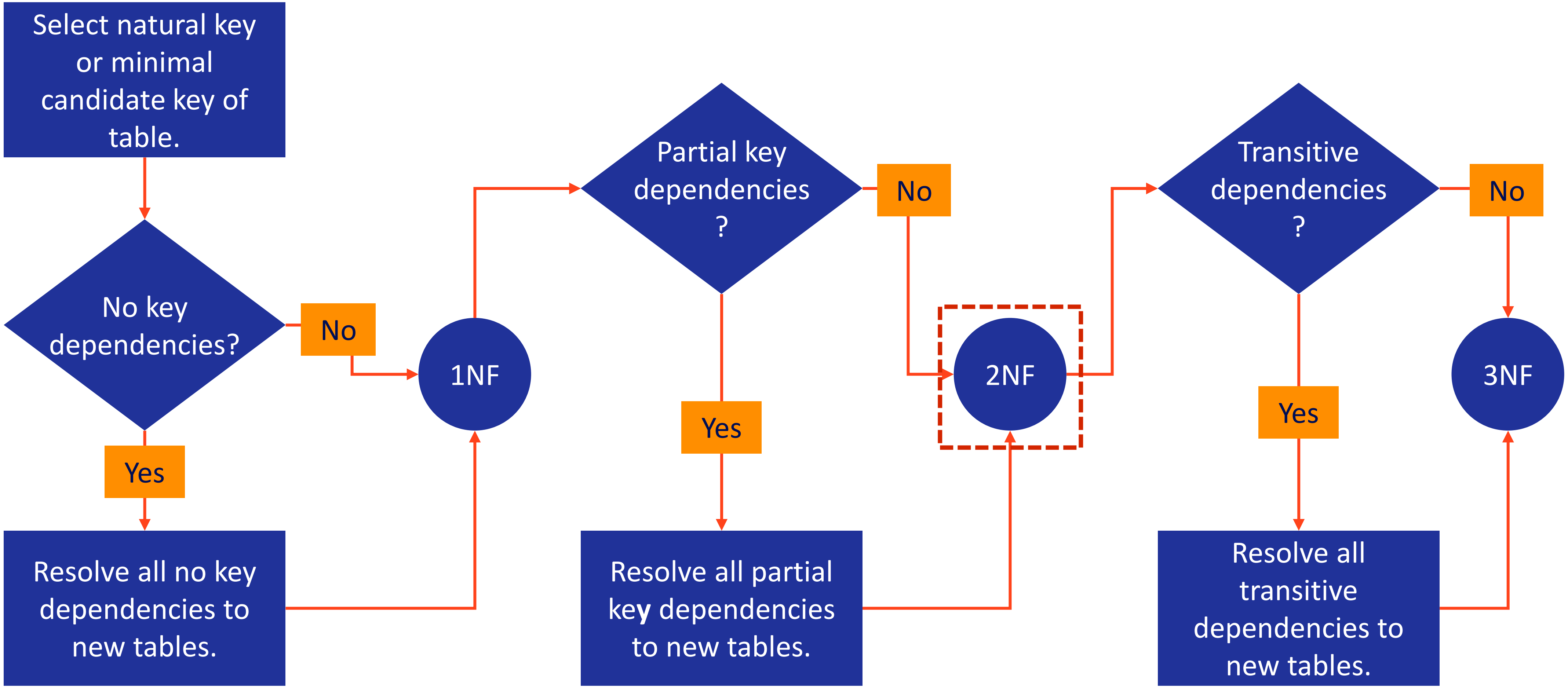
Demo: Resolve Partial Key Dependencies

The End

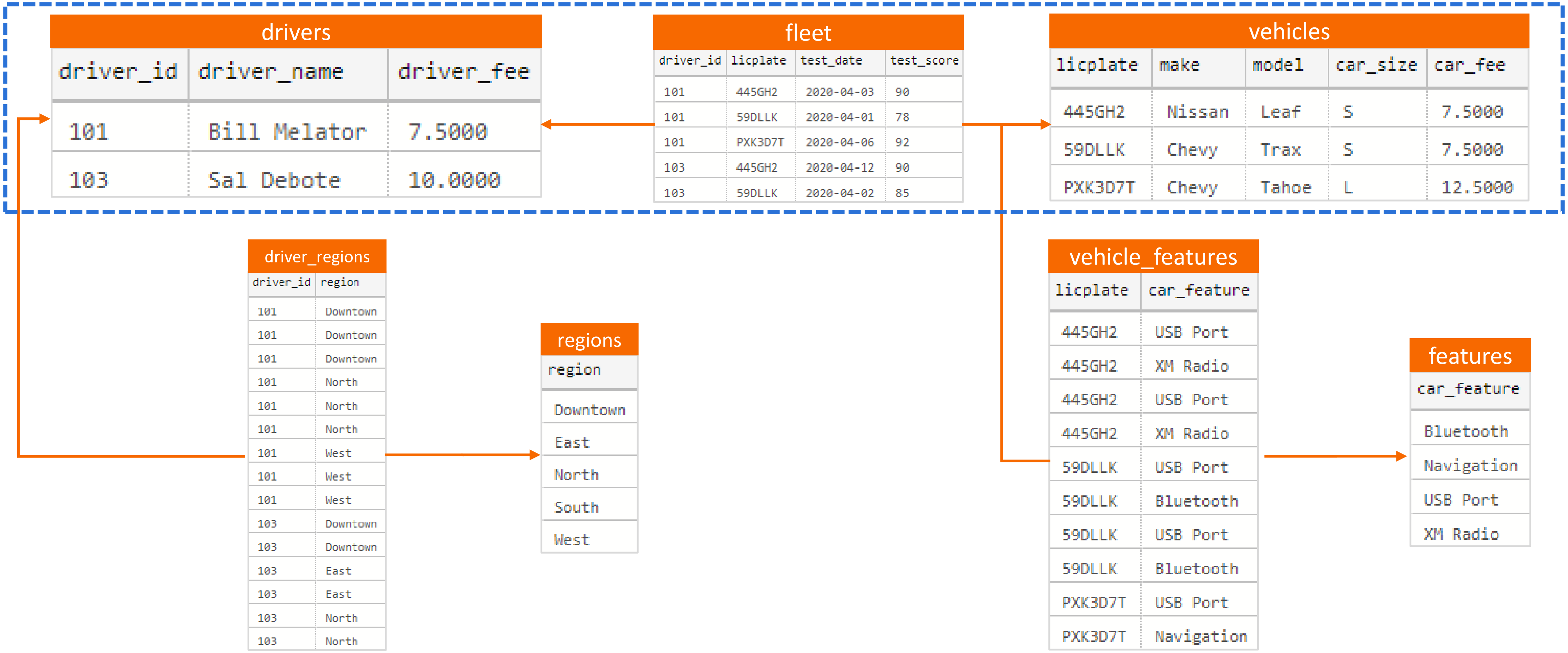
Syracuse University
School of Information Studies

Second Normal Form

Second Normal Form


Select natural key or minimal candidate key of table.
No key dependencies?
No
Yes
Resolve all no key dependencies to new tables.
1NF
Partial key dependencies ?
No
Yes
Resolve all partial key dependencies to new tables.
2NF
Transitive dependencies ?
No
Yes
Resolve all transitive dependencies to new tables.
3NF

Cuse Rides: Resolved Partial Key Dependencies, Now in 2NF



drivers

driver_id	driver_name	driver_fee
101	Bill Melator	7.5000
103	Sal Debote	10.0000

fleet

driver_id	licplate	test_date	test_score
101	445GH2	2020-04-03	90
101	59DLLK	2020-04-01	78
101	PXK3D7T	2020-04-06	92
103	445GH2	2020-04-12	90
103	59DLLK	2020-04-02	85

vehicles

licplate	make	model	car_size	car_fee
445GH2	Nissan	Leaf	S	7.5000
59DLLK	Chevy	Trax	S	7.5000
PXK3D7T	Chevy	Tahoe	L	12.5000

driver_regions

driver_id	region
101	Downtown
101	Downtown
101	Downtown
101	North
101	North
101	North
101	West
101	West
101	West
103	Downtown
103	Downtown
103	East
103	East
103	North
103	North

regions

region
Downtown
East
North
South
West

vehicle_features

licplate	car_feature
445GH2	USB Port
445GH2	XM Radio
445GH2	USB Port
445GH2	XM Radio
59DLLK	USB Port
59DLLK	Bluetooth
59DLLK	USB Port
59DLLK	Bluetooth
PXK3D7T	USB Port
PXK3D7T	Navigation

features

car_feature
Bluetooth
Navigation
USB Port
XM Radio

Syracuse University
School of Information Studies

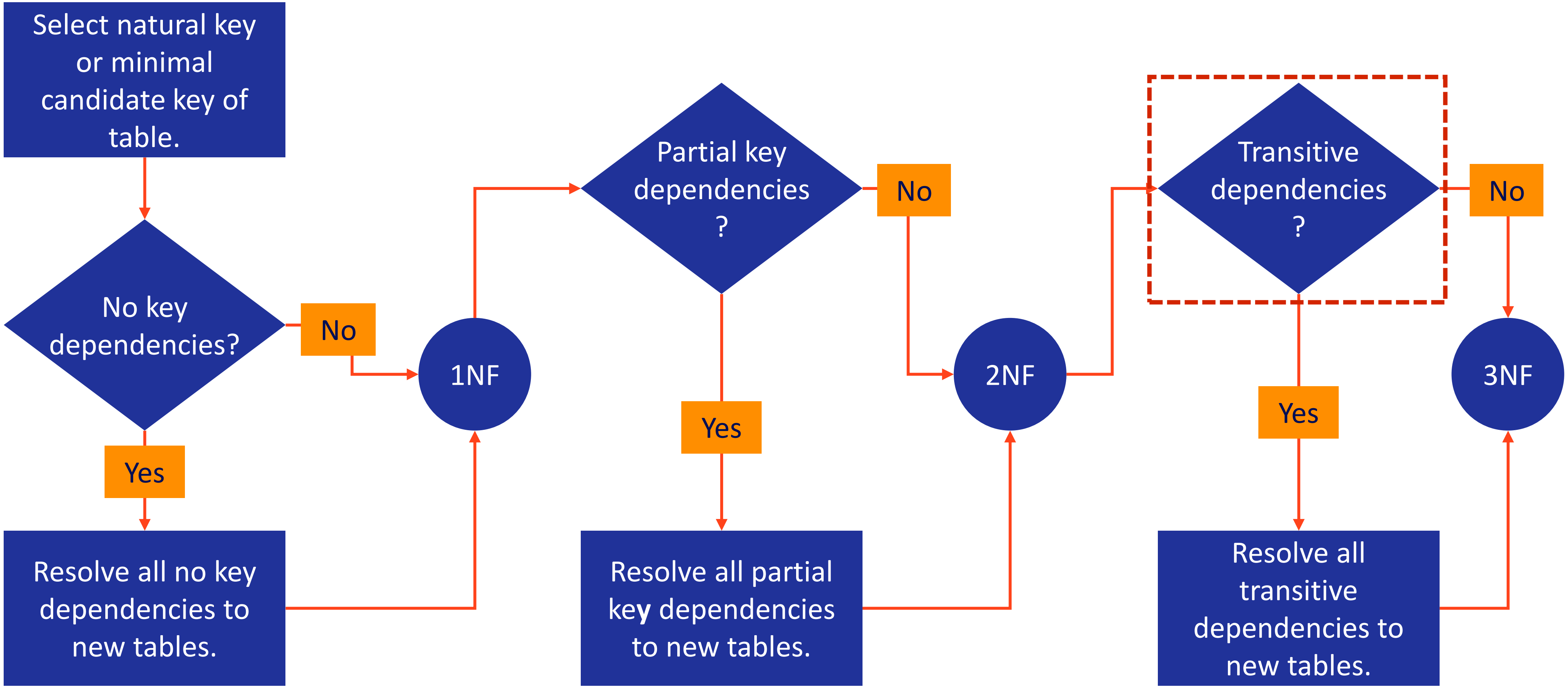
Second Normal Form

The End

Syracuse University
School of Information Studies

Identifying Transitive Dependencies

Identifying Transitive Dependencies



Select natural key or minimal candidate key of table.

No key dependencies?

No

Yes

Resolve all no key dependencies to new tables.

1NF

Partial key dependencies ?

No

Yes

Resolve all partial key dependencies to new tables.

2NF

Transitive dependencies ?

No

Yes

Resolve all transitive dependencies to new tables.

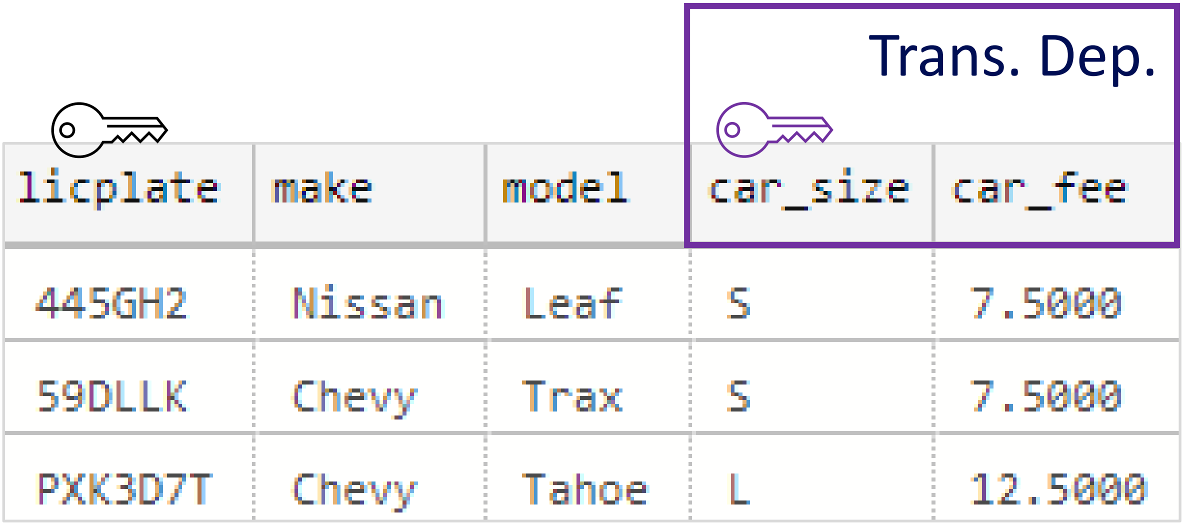
3NF

Identifying Transitive Dependencies (cont.)

- When a no key column functions as a key for one or more other columns, we have a transitive dependency.
- Transitive dependencies happen when a one-to-many relationship is embedded within a table.
- To identify transitive dependencies, we must focus on the no key columns within the table.
- The table requires at least two no key columns for a transitive dependency to exist.

Cuse Rides: Transitive Dependencies

- The Cuse Rides vehicles table a transitive dependency.
- The car_size column acts as a key for values in the car_fee column.
- It is transitive because car_fee does not require the licplate directly.



licplate	make	model	car_size	car_fee
445GH2	Nissan	Leaf	S	7.5000
59DLLK	Chevy	Trax	S	7.5000
PXK3D7T	Chevy	Tahoe	L	12.5000

Syracuse University
School of Information Studies

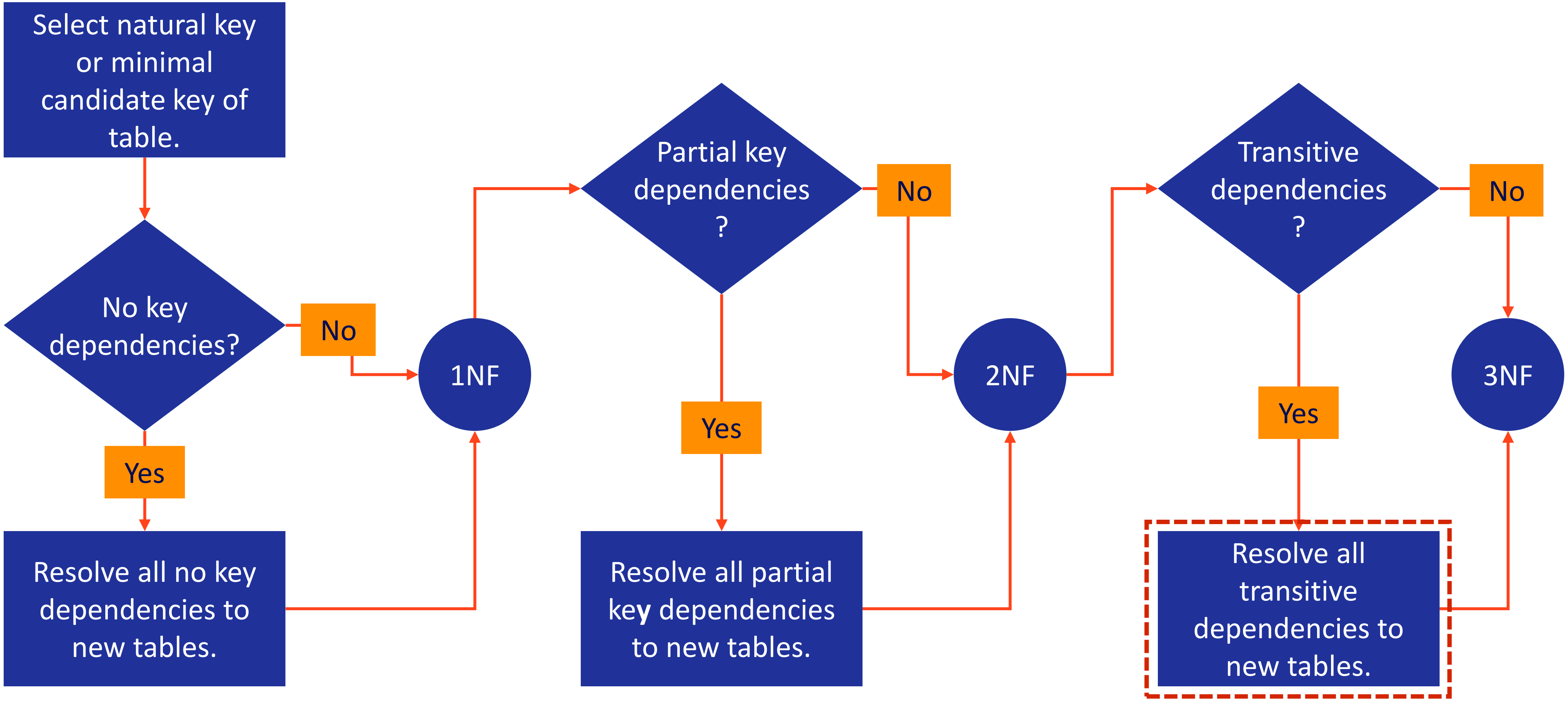
Identifying Transitive Dependencies

The End

Syracuse University
School of Information Studies

Resolving Transitive Dependencies

Resolving Transitive Dependencies



Select natural key or minimal candidate key of table.

No key dependencies?

No

Yes

Resolve all no key dependencies to new tables.

1NF

Partial key dependencies ?

No

Yes

Resolve all partial key dependencies to new tables.

2NF

Transitive dependencies ?

No

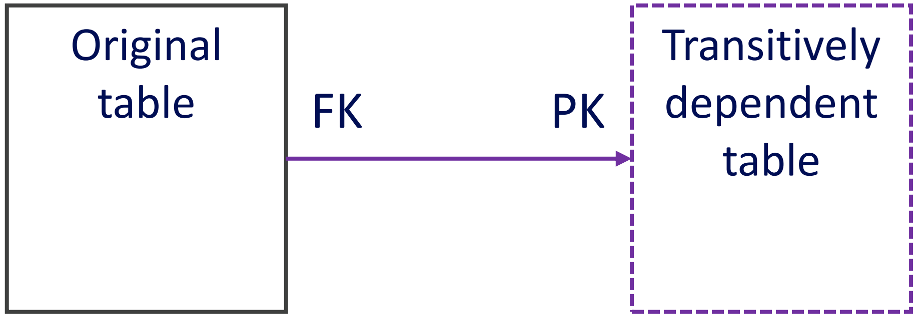
Yes

Resolve all transitive dependencies to new tables.

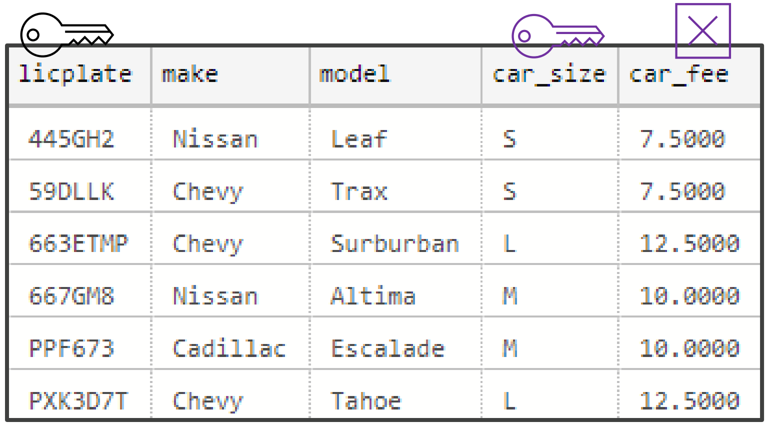
3NF

Resolving Transitive Dependencies (cont.)

- Transitively dependent columns will be removed from the current table and added to a new table.
- The primary key of the new table is the column on which they are transitively dependent, the column acting as key.
- The original table's column acting as key is now the foreign key referencing the new table.



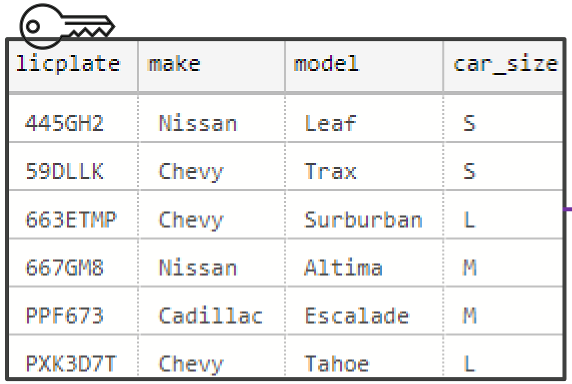
Cuse Rides: Resolve Transitive Dependencies



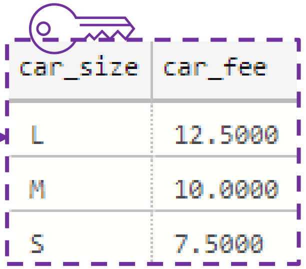
licplate	make	model	car_size	car_fee
445GH2	Nissan	Leaf	S	7.5000
59DLLK	Chevy	Trax	S	7.5000
663ETMP	Chevy	Surburban	L	12.5000
667GM8	Nissan	Altima	M	10.0000
PPF673	Cadillac	Escalade	M	10.0000
PXK3D7T	Chevy	Tahoe	L	12.5000

Vehicles table

Dependencies removed



licplate	make	model	car_size
445GH2	Nissan	Leaf	S
59DLLK	Chevy	Trax	S
663ETMP	Chevy	Surburban	L
667GM8	Nissan	Altima	M
PPF673	Cadillac	Escalade	M
PXK3D7T	Chevy	Tahoe	L



car_size	car_fee
L	12.5000
M	10.0000
S	7.5000

Transitive dependencies for car_size

Syracuse University
School of Information Studies

Resolving Transitive Dependencies

The End

Syracuse University
School of Information Studies

Demo

Resolve Transitive Dependencies

Demo: Resolve Transitive Dependencies

- We will use the Azure Data Studio application.
- We will use the demo database, cr_fleet table.
- Create the tables.
  - cr_vehicles_sizes (transitive table)
- Migrate data.
  - SELECT DISTINCT to reduce rows
  - INSERT INTO SELECT, SELECT INTO

Syracuse University
School of Information Studies

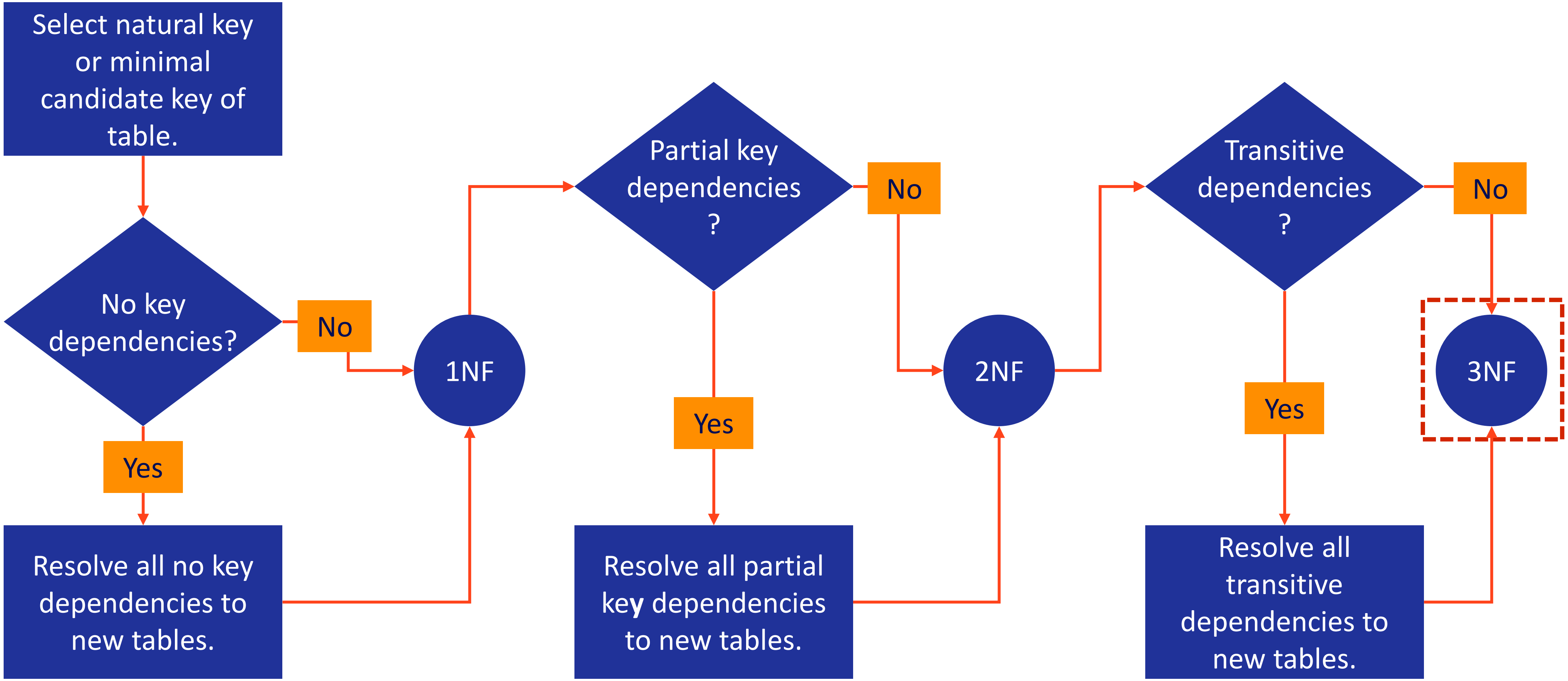
Demo: Resolve Transitive Dependencies

The End

Syracuse University
School of Information Studies

Third Normal Form

Third Normal Form



Select natural key or minimal candidate key of table.

No key dependencies?

No

Yes

Resolve all no key dependencies to new tables.

1NF

Partial key dependencies ?

No

Yes

Resolve all partial key dependencies to new tables.

2NF

Transitive dependencies ?

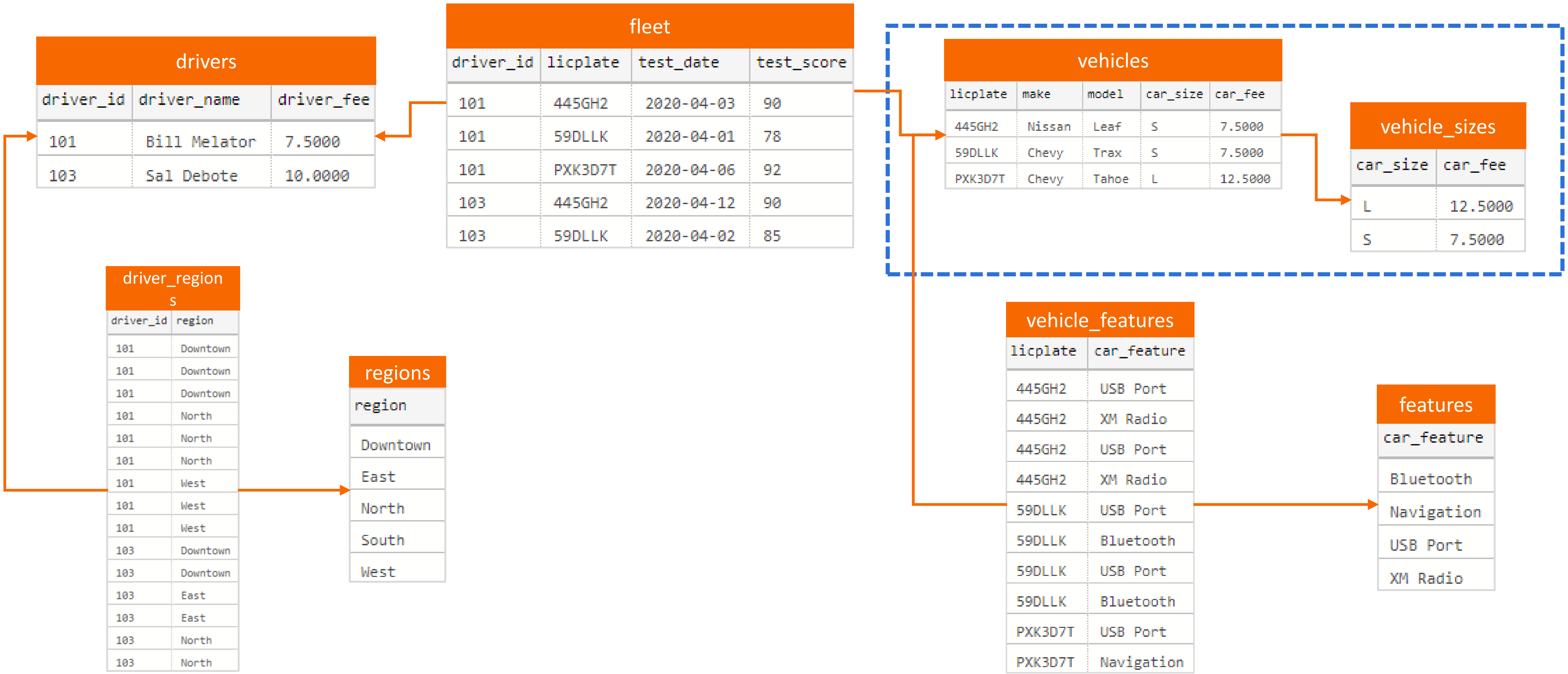
No

Yes

Resolve all transitive dependencies to new tables.

3NF

Cuse Rides: Resolved Transitive Dependencies, Now in 3NF



drivers

driver_id	driver_name	driver_fee
101	Bill Melator	7.5000
103	Sal Debote	10.0000

fleet

driver_id	licplate	test_date	test_score
101	445GH2	2020-04-03	90
101	59DLLK	2020-04-01	78
101	PXK3D7T	2020-04-06	92
103	445GH2	2020-04-12	90
103	59DLLK	2020-04-02	85

vehicles

licplate	make	model	car_size	car_fee
445GH2	Nissan	Leaf	S	7.5000
59DLLK	Chevy	Trax	S	7.5000
PXK3D7T	Chevy	Tahoe	L	12.5000

vehicle_sizes

car_size	car_fee
L	12.5000
S	7.5000

driver_regions

driver_id	region
101	Downtown
101	Downtown
101	Downtown
101	North
101	North
101	North
101	West
101	West
101	West
103	Downtown
103	Downtown
103	East
103	East
103	North
103	North

regions

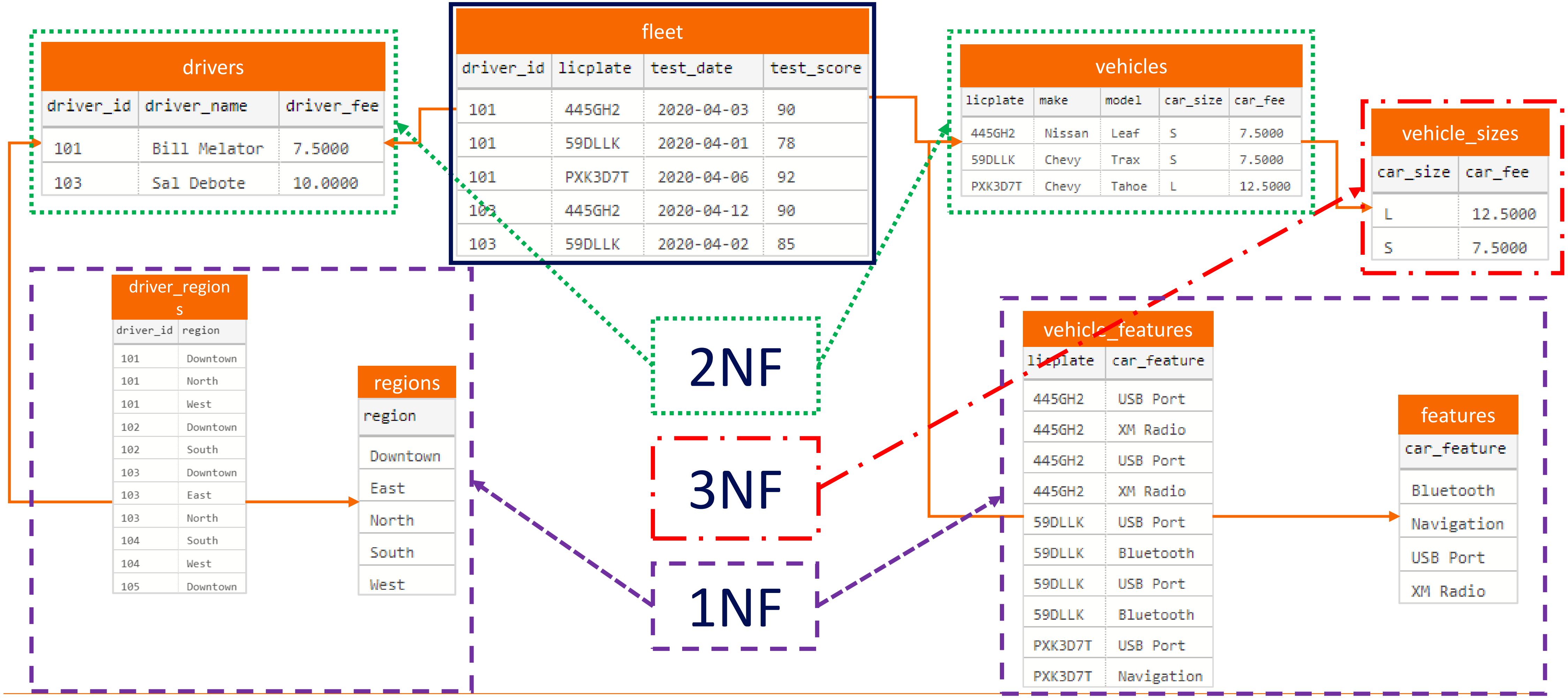
region
Downtown
East
North
South
West

vehicle_features

licplate	car_feature
445GH2	USB Port
445GH2	XM Radio
445GH2	USB Port
445GH2	XM Radio
59DLLK	USB Port
59DLLK	Bluetooth
59DLLK	USB Port
59DLLK	Bluetooth
PXK3D7T	USB Port
PXK3D7T	Navigation

features

car_feature
Bluetooth
Navigation
USB Port
XM Radio



fleet

driver_id	licplate	test_date	test_score
101	445GH2	2020-04-03	90
101	59DLLK	2020-04-01	78
101	PXK3D7T	2020-04-06	92
103	445GH2	2020-04-12	90
103	59DLLK	2020-04-02	85

drivers

driver_id	driver_name	driver_fee
101	Bill Melator	7.5000
103	Sal Debote	10.0000

vehicles

licplate	make	model	car_size	car_fee
445GH2	Nissan	Leaf	S	7.5000
59DLLK	Chevy	Trax	S	7.5000
PXK3D7T	Chevy	Tahoe	L	12.5000

vehicle_sizes

car_size	car_fee
L	12.5000
S	7.5000

driver_regions

driver_id	region
101	Downtown
101	North
101	West
102	Downtown
102	South
103	Downtown
103	East
103	North
104	South
104	West
105	Downtown

regions

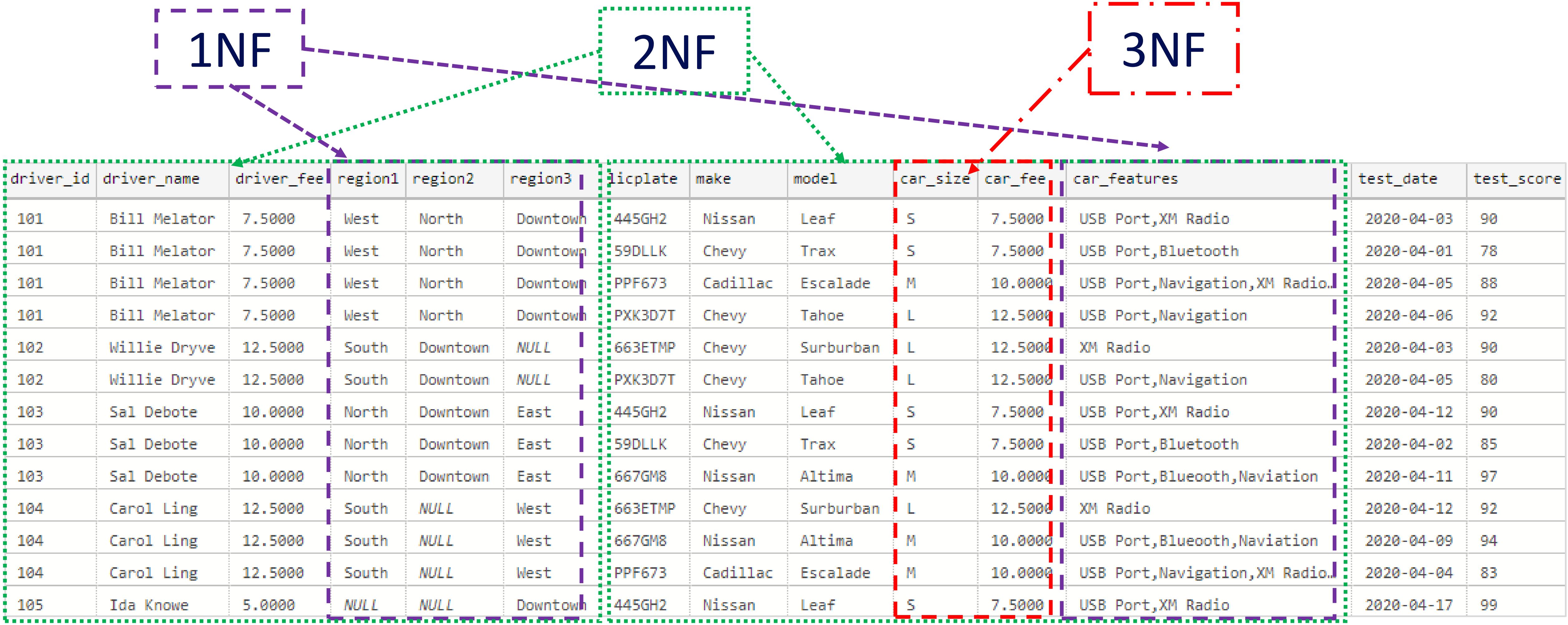
region
Downtown
East
North
South
West

vehicle_features

licplate	car_feature
445GH2	USB Port
445GH2	XM Radio
445GH2	USB Port
445GH2	XM Radio
59DLLK	USB Port
59DLLK	Bluetooth
59DLLK	USB Port
59DLLK	Bluetooth
PXK3D7T	USB Port
PXK3D7T	Navigation

features

car_feature
Bluetooth
Navigation
USB Port
XM Radio



driver_id	driver_name	driver_fee	region1	region2	region3	licplate	make	model	car_size	car_fee	car_features	test_date	test_score
101	Bill Melator	7.5000	West	North	Downtown	445GH2	Nissan	Leaf	S	7.5000	USB Port,XM Radio	2020-04-03	90
101	Bill Melator	7.5000	West	North	Downtown	59DLLK	Chevy	Trax	S	7.5000	USB Port,Bluetooth	2020-04-01	78
101	Bill Melator	7.5000	West	North	Downtown	PPF673	Cadillac	Escalade	M	10.0000	USB Port,Navigation,XM Radio..	2020-04-05	88
101	Bill Melator	7.5000	West	North	Downtown	PXK3D7T	Chevy	Tahoe	L	12.5000	USB Port,Navigation	2020-04-06	92
102	Willie Dryve	12.5000	South	Downtown	NULL	663ETMP	Chevy	Surburban	L	12.5000	XM Radio	2020-04-03	90
102	Willie Dryve	12.5000	South	Downtown	NULL	PXK3D7T	Chevy	Tahoe	L	12.5000	USB Port,Navigation	2020-04-05	80
103	Sal Debote	10.0000	North	Downtown	East	445GH2	Nissan	Leaf	S	7.5000	USB Port,XM Radio	2020-04-12	90
103	Sal Debote	10.0000	North	Downtown	East	59DLLK	Chevy	Trax	S	7.5000	USB Port,Bluetooth	2020-04-02	85
103	Sal Debote	10.0000	North	Downtown	East	667GM8	Nissan	Altima	M	10.0000	USB Port,Blueooth,Naviation	2020-04-11	97
104	Carol Ling	12.5000	South	NULL	West	663ETMP	Chevy	Surburban	L	12.5000	XM Radio	2020-04-12	92
104	Carol Ling	12.5000	South	NULL	West	667GM8	Nissan	Altima	M	10.0000	USB Port,Blueooth,Naviation	2020-04-09	94
104	Carol Ling	12.5000	South	NULL	West	PPF673	Cadillac	Escalade	M	10.0000	USB Port,Navigation,XM Radio..	2020-04-04	83
105	Ida Knowe	5.0000	NULL	NULL	Downtown	445GH2	Nissan	Leaf	S	7.5000	USB Port,XM Radio	2020-04-17	99

Syracuse University
School of Information Studies

Third Normal Form

The End

Syracuse University
School of Information Studies

Demo

Add Foreign Keys

Demo: Add Foreign Keys

- We will use the Azure Data Studio application.
- We will use the demo database
- Now that we are in 3NF, we can add the FKs to all the new tables.

Syracuse University
School of Information Studies

Demo: Add Foreign Keys

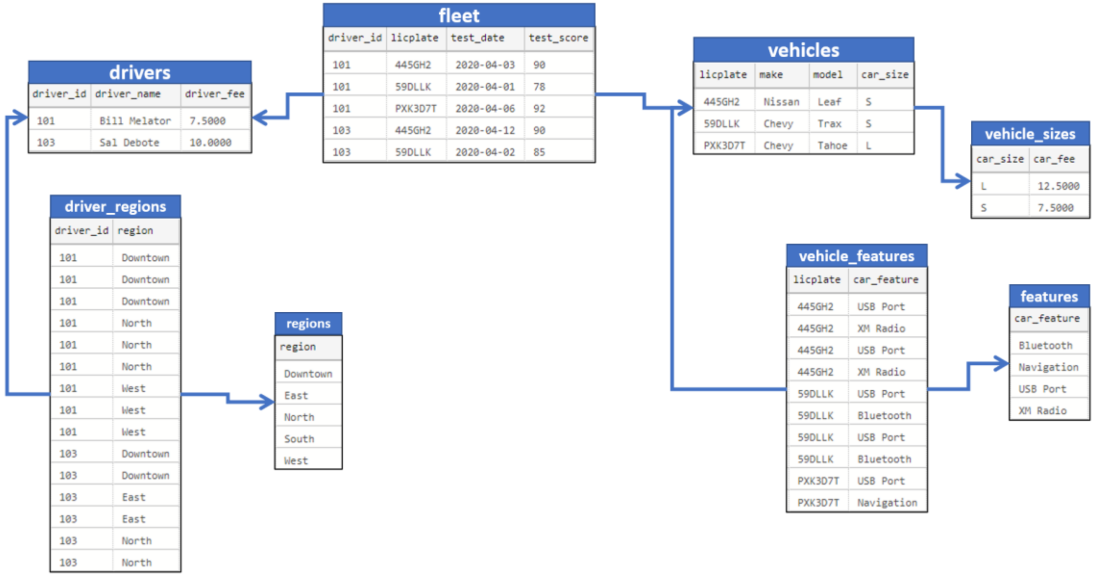
The End

Syracuse University
School of Information Studies

Denormalization

Are We Taking It Too Far?

- We went from one table…
- To eight tables!
- Yes, the design is better, but is it necessary?



drivers

driver_id	driver_name	driver_fee
101	Bill Melator	7.5000
103	Sal Debote	10.0000

fleet

driver_id	licplate	test_date	test_score
101	445GH2	2020-04-03	90
101	59DLLK	2020-04-01	78
101	PXK3D7T	2020-04-06	92
103	445GH2	2020-04-12	90
103	59DLLK	2020-04-02	85

vehicles

licplate	make	model	car_size
445GH2	Nissan	Leaf	S
59DLLK	Chevy	Trax	S
PXK3D7T	Chevy	Tahoe	L

vehicle_sizes

car_size	car_fee
L	12.5000
S	7.5000

driver_regions

driver_id	region
101	Downtown
101	Downtown
101	Downtown
101	North
101	North
101	North
101	West
101	West
101	West
103	Downtown
103	Downtown
103	East
103	East
103	North
103	North

regions

region
Downtown
East
North
South
West

vehicle_features

licplate	car_feature
445GH2	USB Port
445GH2	XM Radio
445GH2	USB Port
445GH2	XM Radio
59DLLK	USB Port
59DLLK	Bluetooth
59DLLK	USB Port
59DLLK	Bluetooth
PXK3D7T	USB Port
PXK3D7T	Navigation

features

car_feature
Bluetooth
Navigation
USB Port
XM Radio

fleet

driver_id	driver_name	driver_fee	region1	region2	region3	licplate	make	model	car_size	car_fee	car_features	test_date	test_score
101	Bill Melator	7.5000	West	North	Downtown	445GH2	Nissan	Leaf	S	7.5000	USB Port,XM Radio	2020-04-03	90
101	Bill Melator	7.5000	West	North	Downtown	59DLLK	Chevy	Trax	S	7.5000	USB Port,Bluetooth	2020-04-01	78
101	Bill Melator	7.5000	West	North	Downtown	PXK3D7T	Chevy	Tahoe	L	12.5000	USB Port,Navigation	2020-04-06	92
103	Sal Debote	10.0000	North	Downtown	East	445GH2	Nissan	Leaf	S	7.5000	USB Port,XM Radio	2020-04-12	90
103	Sal Debote	10.0000	North	Downtown	East	59DLLK	Chevy	Trax	S	7.5000	USB Port,Bluetooth	2020-04-02	85

How Far Should We Normalize?

- It is a trade-off
    - Higher normal forms are complex but introduce fewer chances for data anomalies.
- General rules
    - Is 1NF required? It depends. 1NF introduces two new tables and quite some complexity but makes the data easier to query and maintain.
    - Is 2NF required? Yes. You are hiding an M-M relationship.
    - Is 3NF required? It depends. Maybe? Is the table used elsewhere like for zip code or address?
- Never design or normalize in a vacuum, following the rules blindly.
- Let the needs of the database application dictate the design!

Syracuse University
School of Information Studies

Denormalization

The End

Syracuse University
School of Information Studies

Summary

Summary

- Data normalization improves our table designs by eliminating data anomalies that can cause redundancies and inconsistencies.
- Normalization is a series of checks based on columns with no key dependencies, partial key dependencies, or transitive dependencies.
- The normalization process involves identifying and then resolving each of these types of dependencies.
- Normalization is not a cure-all; the database application and its use should be considered.

Syracuse University
School of Information Studies

Summary

The End