



Week 8

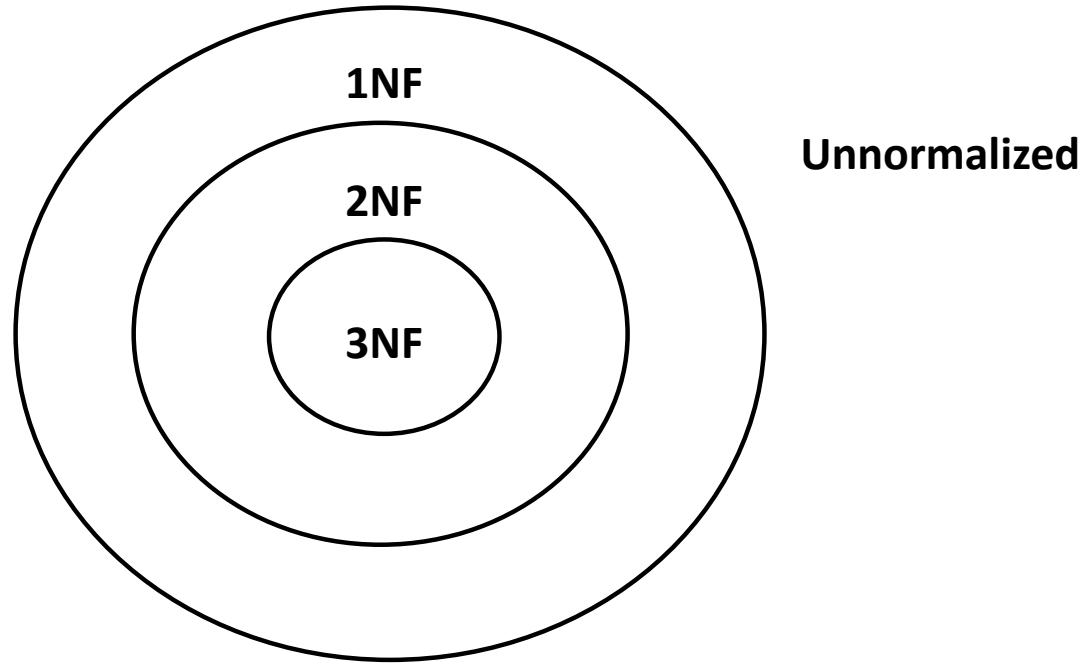
Foundations of Data Warehousing

October 12, 2026

Review Milestone 5 (data cleaning worksheet)



Normalization Concepts



Unnormalized to 1NF

Rule: A database schema is in 1NF iff all attributes have scalar values.

Student_Semester

<u>EID</u>	<u>Semester</u>	Classes		
alice1	Fall26	Stats	A	Lee
		DB	A	Patel
		Alg	A-	Garcia
bob20	Fall26	DB	A	Patel
		Alg	B+	Garcia
carol30	Fall26	Stats	A-	Lee
		Alg	B+	Garcia

Unnormalized



Student_Semester'

<u>EID</u>	<u>Semester</u>	<u>Class</u>	<u>Grade</u>	<u>Instructor</u>
alice1	Fall26	Stats	A	Lee
alice1	Fall26	DB	A	Patel
alice1	Fall26	Alg	A-	Garcia
bob20	Fall26	DB	A	Patel
bob20	Fall26	Alg	B+	Garcia
carol30	Fall26	Stats	A-	Lee
carol30	Fall26	Alg	B+	Garcia

1NF

Functional Dependencies

Definition:

If two records agree on the attributes

$A_1, A_2, \dots, A_n$

then they must also agree on the attributes

$B_1, B_2, \dots, B_n$

Formally:

$A_1, A_2, \dots, A_n \rightarrow B_1, B_2, \dots, B_n$

1NF to 2NF

Rule: A database schema is in 2NF iff it is in 1NF and there exists no partial FDs on the primary key (i.e. all non-key attributes must be dependent on the entire PK)

Student_Semester'

<u>EID</u>	<u>Semester</u>	<u>Class</u>	<u>Grade</u>	<u>Instructor</u>
alice1	Fall26	Stats	A	Lee
alice1	Fall26	DB	A	Patel
alice1	Fall26	Alg	A-	Garcia
bob20	Fall26	DB	A	Patel
bob20	Fall26	Alg	B+	Garcia
carol30	Fall26	Stats	A-	Lee
carol30	Fall26	Alg	B+	Garcia

1NF

Enrollment

<u>EID</u>	<u>Semester</u>	<u>Class</u>	<u>Grade</u>
alice1	Fall26	Stats	A
alice1	Fall26	DB	A
alice1	Fall26	Alg	A-
bob20	Fall26	DB	A
bob20	Fall26	Alg	B+
carol30	Fall26	Stats	A-
carol30	Fall26	Alg	B+

2NF

Class_Offering

<u>Semester</u>	<u>Class</u>	<u>Instructor</u>
Fall26	Stats	Lee
Fall26	DB	Patel
Fall26	Alg	Garcia

2NF

2NF to 3NF

Rule: A database schema is in 3NF iff it is in 2NF and there exists no non-key attributes that are functionally determined by other non-key attributes.

Student_Major

2NF

<u>EID</u>	Name	Major	College
alice1	Alice	Math	Natural Sciences
bob20	Bob	CS	Natural Sciences
carol30	Carol	Math	Natural Sciences

Student_Major'

3NF

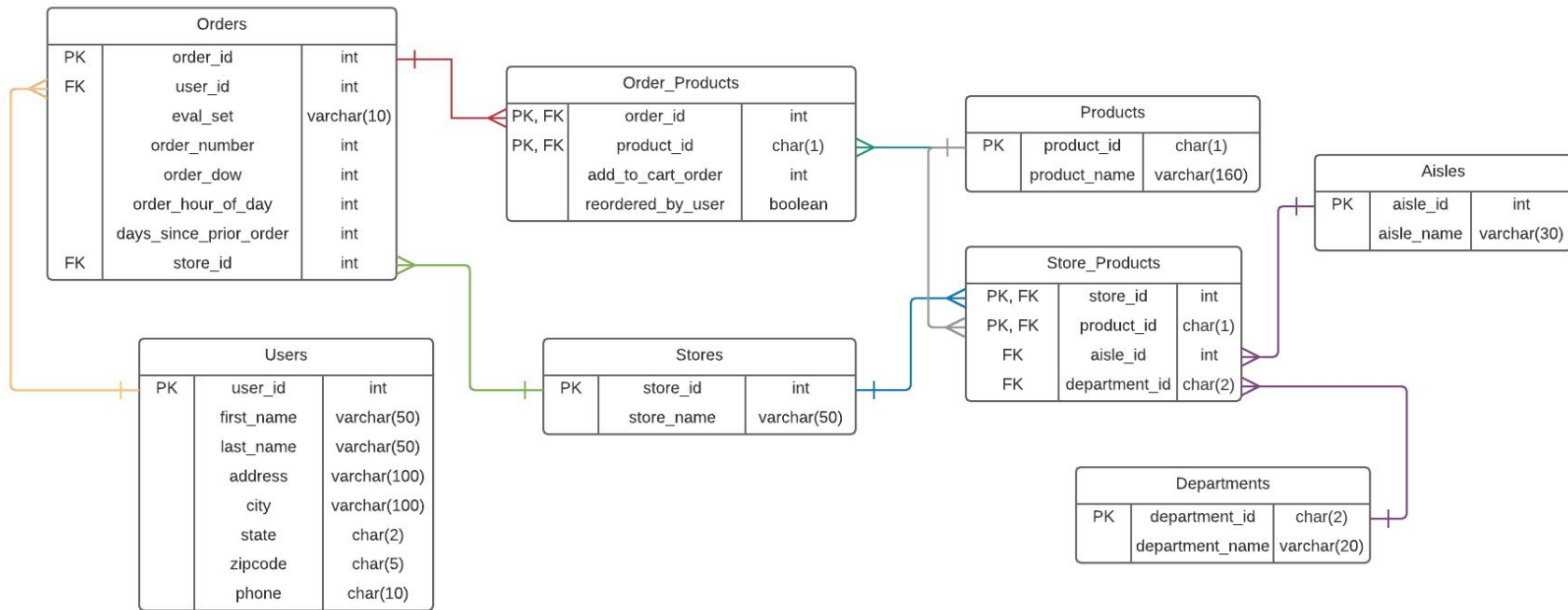
<u>EID</u>	Name	Major
alice1	Alice	Math
bob20	Bob	CS
carol30	Carol	Math

Major_College

3NF

<u>Major</u>	College
Math	Natural Sciences
CS	Natural Sciences

Exercise: What is the highest normal form of the Instacart schema?



See [sample data](#) for each table.

Exercise: What is the highest normal form of the Instacart schema?

