# Database Design

Monday, January 23, 2017

# Agenda

- Announcements
- Reading Quiz
- Basic Concepts
- Data Anomalies
- Practice Problem #1
- Conceptual Design

# Announcements

- Create github account
- Form teams and send email by Wed:
  - Team members full names, EIDs, Github usernames
  - Team name
  - Use email subject line: CS327E Team XYZ, where XYZ is your team name
  - Send email to me and both TAs, copy your partner on the email
- Lowest 2 quizzes and participation scores will be dropped

# Q1: It is worth the time and effort needed to design your database well from the start.

a) True

b) False

# Q2: Poor database design can lead to...

a) Data inconsistency

b) Unnecessary duplicated data

c) Retrieving Inaccurate Data

d) Untimely queries

e) All of the above

# Q3: Which is not a data relationship?

a) One-to-one

b) One-to-many

c) Many-to-many

d) Strong-to-weak

# Q4: A primary key should use arbitrary identifiers or concatenation of arbitrary identifiers whenever possible.

a) True

b) False

# Q5: Which is not a normal form?

a) First normal form

b) Second normal form

c) Third normal form

d) Seventh normal form

e) Boyce–Codd Normal Form

# Basic Terminology

- Entity / Relation / Table
- Column / Field / Attribute
- Record / Row / Tuple

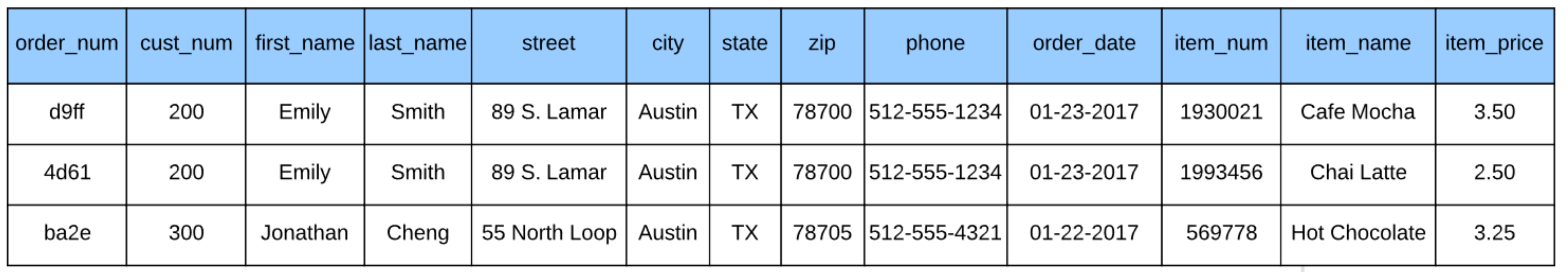
# Basic Concepts

- Schema
- Keys
- Relationships
- Constraints and data integrity
- Data consistency and normal forms
- SQL (DDL and DML)

# Data Anomalies

- Insertion anomaly
- Update anomaly
- Deletion anomaly

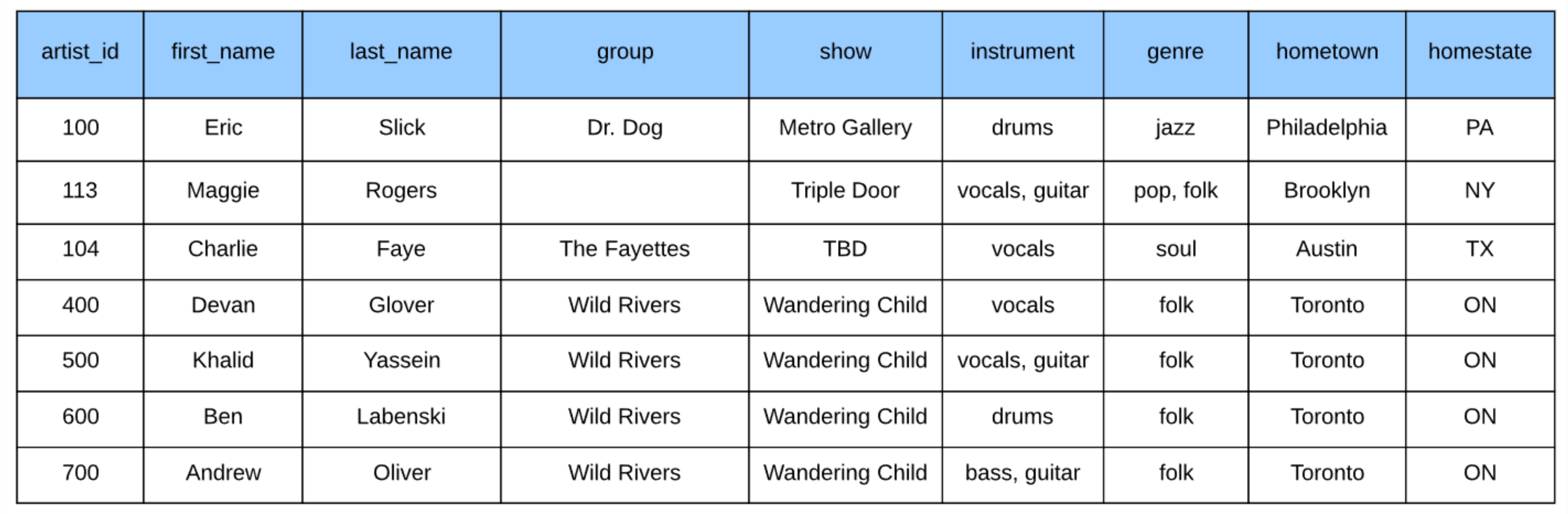
Orders(order_num, cust_num, first_name, last_name, city, state, zip, phone, order_date, item_num, item_name, item_price)



| order_num | cust_num | first_name | last_name | street | city | state | zip | phone | order_date | item_num | item_name | item_price |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| d9ff | 200 | Emily | Smith | 89 S. Lamar | Austin | TX | 78700 | 512-555-1234 | 01-23-2017 | 1930021 | Cafe Mocha | 3.50 |
| 4d61 | 200 | Emily | Smith | 89 S. Lamar | Austin | TX | 78700 | 512-555-1234 | 01-23-2017 | 1993456 | Chai Latte | 2.50 |
| ba2e | 300 | Jonathan | Cheng | 55 North Loop | Austin | TX | 78705 | 512-555-4321 | 01-22-2017 | 569778 | Hot Chocolate | 3.25 |

# Practice Problem 1: What’s wrong with this schema? Find instances of insertion, update, and deletion anomalies

Musicians(artist_id, first_name, last_name, group, show, instrument, genre, hometown, homestate)



| artist_id | first_name | last_name | group | show | instrument | genre | hometown | homestate |
|---|---|---|---|---|---|---|---|---|
| 100 | Eric | Slick | Dr. Dog | Metro Gallery | drums | jazz | Philadelphia | PA |
| 113 | Maggie | Rogers |  | Triple Door | vocals, guitar | pop, folk | Brooklyn | NY |
| 104 | Charlie | Faye | The Fayettes | TBD | vocals | soul | Austin | TX |
| 400 | Devan | Glover | Wild Rivers | Wandering Child | vocals | folk | Toronto | ON |
| 500 | Khalid | Yassein | Wild Rivers | Wandering Child | vocals, guitar | folk | Toronto | ON |
| 600 | Ben | Labenski | Wild Rivers | Wandering Child | drums | folk | Toronto | ON |
| 700 | Andrew | Oliver | Wild Rivers | Wandering Child | bass, guitar | folk | Toronto | ON |

# Practice Problem 1

a) None

b) 1-3 data anomalies

c) 4 data anomalies

d) 5 data anomalies

e) > 5 data anomalies

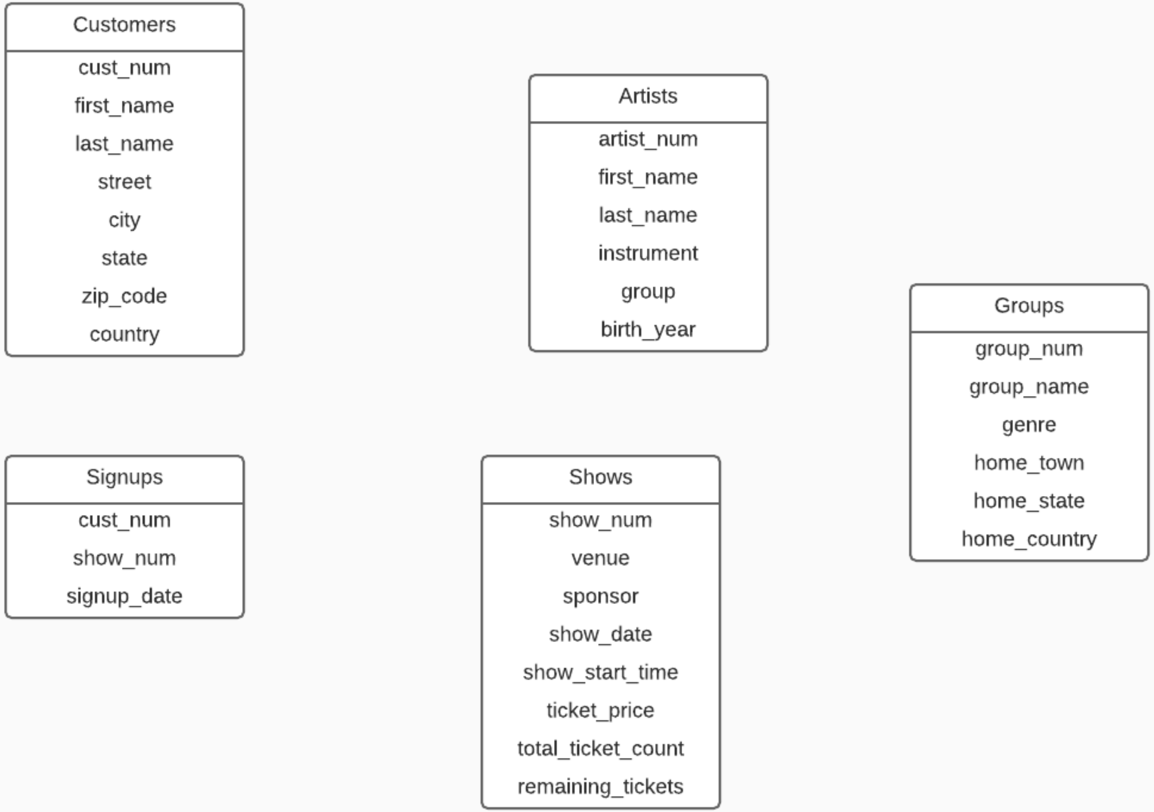
# Lucidchart: Diagramming tool

Steps to sign-up for a Lucidchart account and receive an educational upgrade: https://github.com/wolfier/CS327E/wiki/Setting-up-Lucidchart

# Scenario: SXSW Database

Design a database for the organizers of the music festival to help them gain more insight into their current customer base. Want to use this database answer questions such as **which shows were well-attended last year? Who were the most popular artists and groups/bands based on number of signups? Which customers are loyal and return to the festival year-after-year? Which customers are new and which ones only go to free concerts? Which customers are branching out and sign up for other types of events (networking, etc.)?** These are just a few questions, I’m sure you can think of others to help the organizers of the festival :))

# SXSW Entities



Customers

cust_num
first_name
last_name
street
city
state
zip_code
country

Artists

artist_num
first_name
last_name
instrument
group
birth_year

Groups

group_num
group_name
genre
home_town
home_state
home_country

Signups

cust_num
show_num
signup_date

Shows

show_num
venue
sponsor
show_date
show_start_time
ticket_price
total_ticket_count
remaining_tickets