# Logistics

## Can’t do anything about waitlist

## Names + Grade + Major + Favorite Food

# SQL

## General syntax

### SELECT \_\_\_ FROM \_\_\_\_\_

#### Aliases

#### COUNT

#### AVG

#### MAX

### WHERE \_\_\_\_\_;

#### LIKE

### DISTINCT

### ORDER BY

### GROUP BY

# External Algorithms

## Big Picture

### Motivation: Some way to deal with lots of data

### How it ties in: SQL ( … ) Data

### Approximations:

### Define “runs”, “buffer”, “pages”, “disk”,

#### “IOs” = “Reads”

## Hash

### Algorithm

#### Stream in from disk (B-1 buckets)

#### Form Partition -> write to disk

#### Read Partition -> Form hash table

#### Write out hash table, continue

### 2 Passes: B(B-1)

### Recursive: B(B-1)^(n-1)

## Sort

### K-Way:

#### Stream through to create runs

#### For each B-1, merge and flush out

##### Continue until 1 run

### Runtime:

#### 1 + [log­­­(B-1) [N/B]]

### 2 Passes