Analyzing Food-borne Outbreaks around the release of FDA's 2017 Food Code

Marisa Wong

INFO 521 – Final Presentation

April 21st, 2023

New 2017 Food Code Section
Bandages, Finger Cots, or Finger Stalls

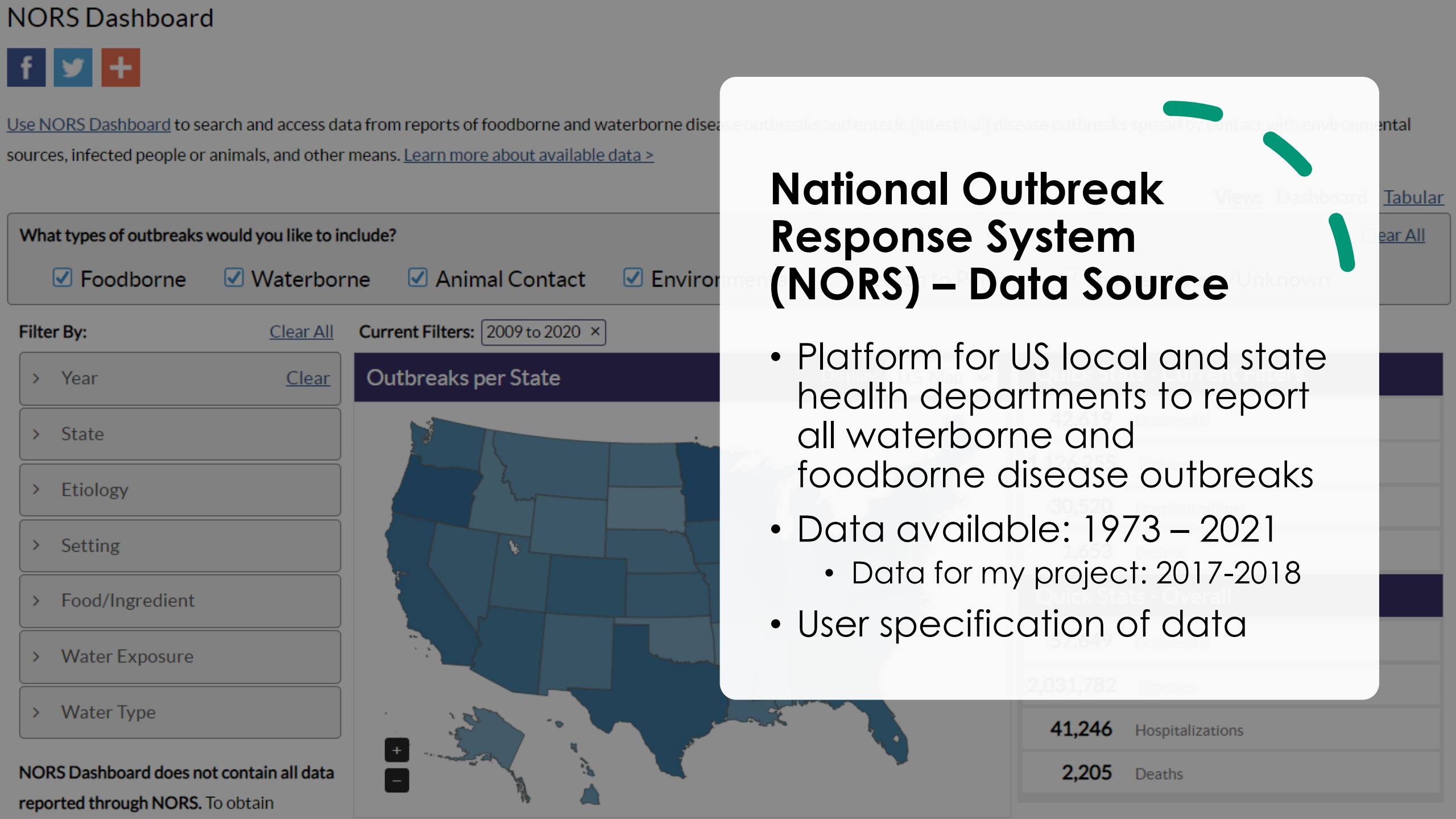
FDA U.S. FOOD & DRUG ADMINISTRATION
NATIONAL RETAIL FOOD TEAM

In the 2017 Food Code, the FDA included a new section stating that single-use gloves shall be worn over bandages, finger cots, or finger stalls, to prevent these items from falling into food during preparation and becoming a physical hazard. Adopting this provision into your regulatory framework can provide effective controls as a means of reducing contamination hazards within retail food establishments, ultimately protecting consumers and industry.

FDA Model Food Code

Provides guidance/framework for states to formulate their own food safety regulations

- Conference for Food Protection (1993)
- Food Sanitation, food and beverage vending, retail food store sanitation
- Adoption methods: short form (by reference) or long form (section-by-section)
- 2017 major update



National Outbreak Response System (NORS) – Data Source

- Platform for US local and state health departments to report all waterborne and foodborne disease outbreaks
- Data available: 1973 – 2021
  - Data for my project: 2017-2018
- User specification of data

MySQL Data Management

SCHEMA: nors_outbreaks

TABLE: timeframe

Column	*Data Type*
outbreak_id*	INT (AI, NN)
outbreak_year	YEAR
outbreak_month	INT
state	VARCHAR(50)

TABLE: etiology

Column	*Data Type*
etiology_id*	INT (AI, NN)
transmission_ mode	VARCHAR(1000)
serotype	VARCHAR(1000)
pathogen	VARCHAR(1000)
etiology_status	VARCHAR(1000)
outbreak_id**	INT (NN)

Legend:
* - Primary key
** - Foreign key
AI – autoincrement
NN – Not Null

Records in original and imported datasets: **8,276**

TABLE: outcomes

Column	*Data Type*
outcomes_id*	INT (AI, NN)
num_illnesses	INT
num_hospitalizations	INT
num_deaths	INT
outbreak_id**	INT (NN)

TABLE: situation

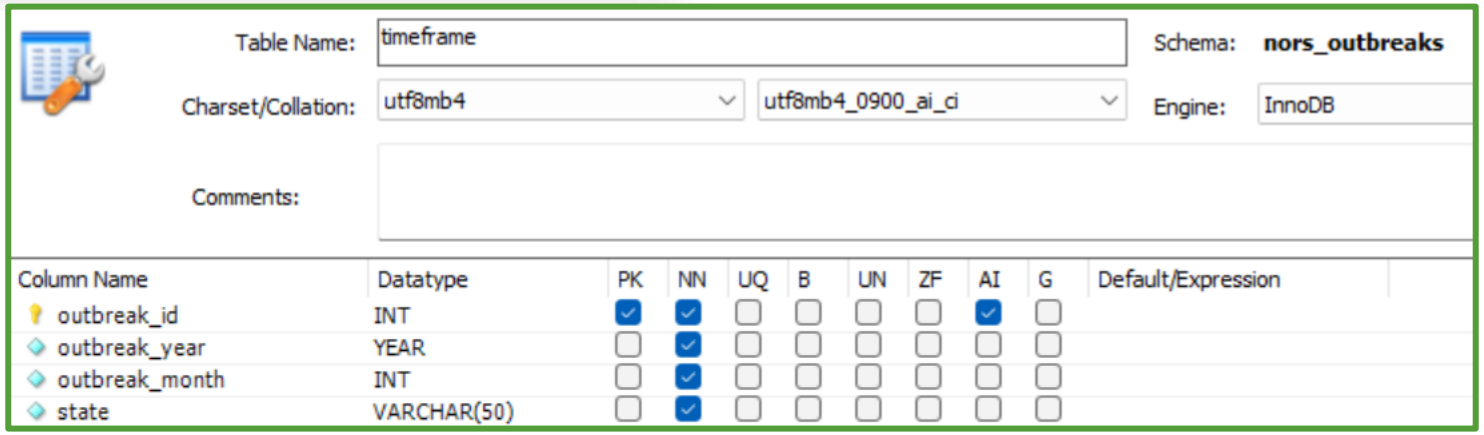
Column	*Data Type*
situation_id*	INT (AI, NN)
setting	VARCHAR(1000)
food_vehicle	VARCHAR(1000)
contaminated_ingredient	VARCHAR(1000)
outbreak_id**	INT (NN)

Creating the Database

```
/* Creating a database with a new database schema */
CREATE SCHEMA nors_outbreaks;
```

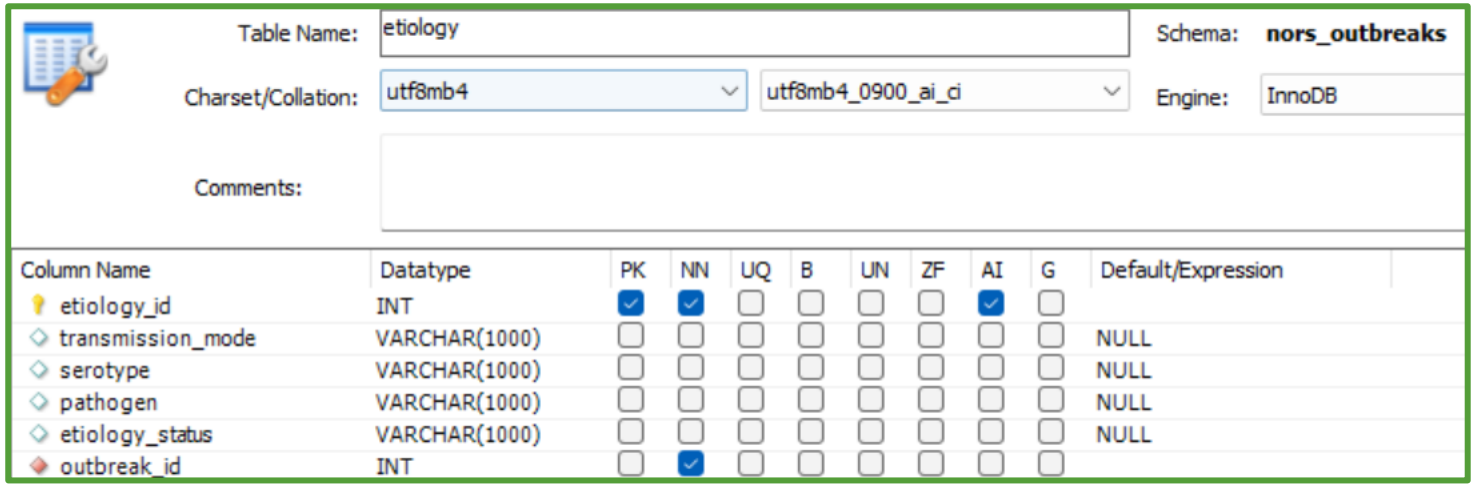
timeframe

```
CREATE TABLE `nors_outbreaks`.`timeframe`(
    `outbreak_id` INT NOT NULL AUTO_INCREMENT,
    `outbreak_year` YEAR NOT NULL,
    `outbreak_month` INT NOT NULL,
    `state` VARCHAR(50) NOT NULL,
    PRIMARY KEY (`outbreak_id`));
```



etiology

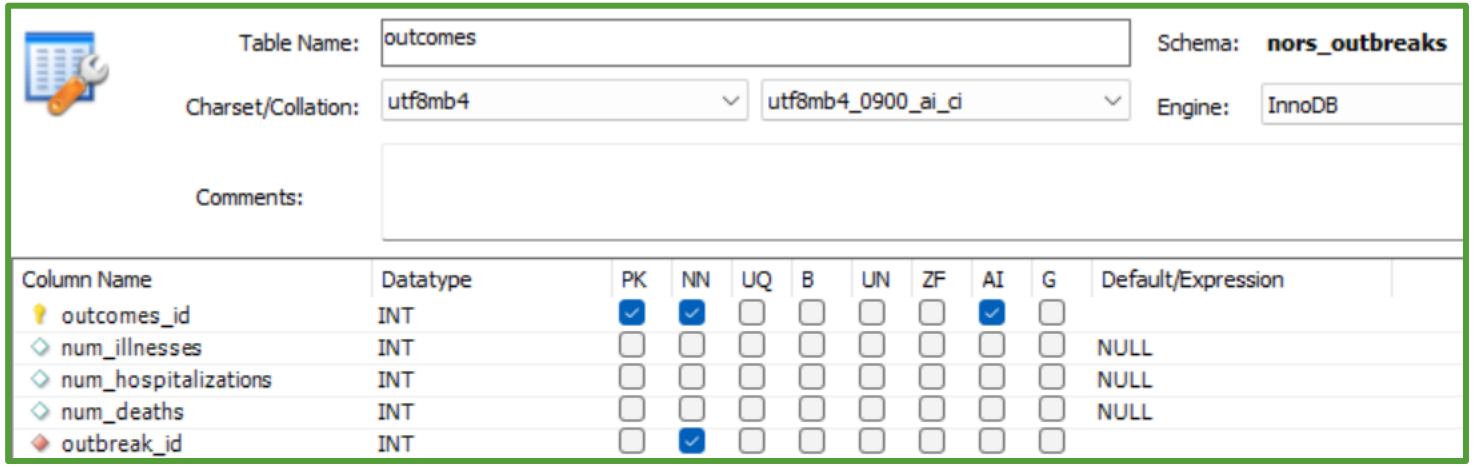
```
CREATE TABLE `nors_outbreaks`.`etiology` (
  `etiology_id` INT NOT NULL AUTO_INCREMENT,
  `transmission_mode` VARCHAR(1000) NULL,
  `serotype` VARCHAR(1000) NULL,
  `pathogen` VARCHAR(1000) NULL,
  `etiology_status` VARCHAR(1000) NULL,
  `outbreak_id` INT NOT NULL,
  PRIMARY KEY (`etiology_id`));
```



Creating the Database

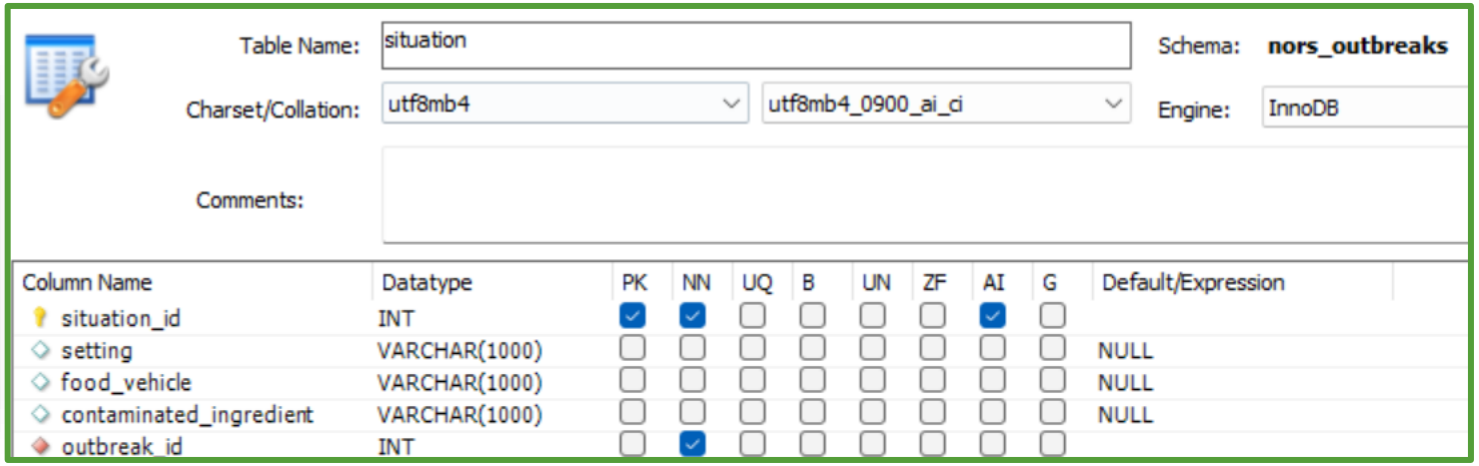
outcomes

```
CREATE TABLE `nors_outbreaks`.`outcomes` (
    `outcomes_id` INT NOT NULL AUTO_INCREMENT,
    `num_illnesses` INT NULL,
    `num_hospitalizations` INT NULL,
    `num_deaths` INT NULL,
    `outbreak_id` INT NOT NULL,
    PRIMARY KEY (`outcomes_id`));
```

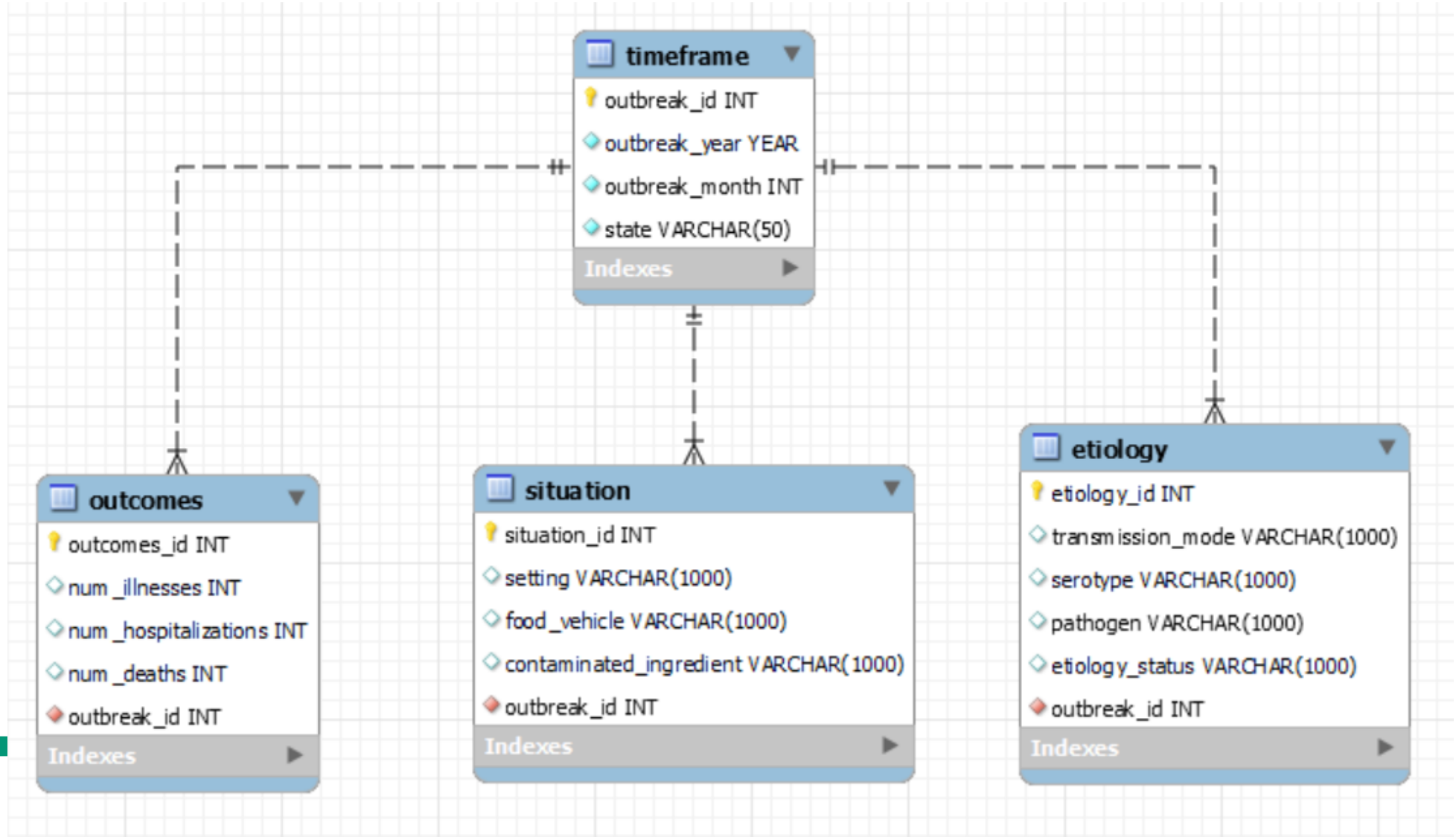


situation

```
CREATE TABLE `nors_outbreaks`.`situation` (
    `situation_id` INT NOT NULL AUTO_INCREMENT,
    `setting` VARCHAR(1000) NULL,
    `food_vehicle` VARCHAR(1000) NULL,
    `contaminated_ingredient` VARCHAR(1000) NULL,
    `outbreak_id` INT NOT NULL,
    PRIMARY KEY (`situation_id`));
```



Entity Relationship Diagram (ERD)



Analytical Research Question & Expected Results

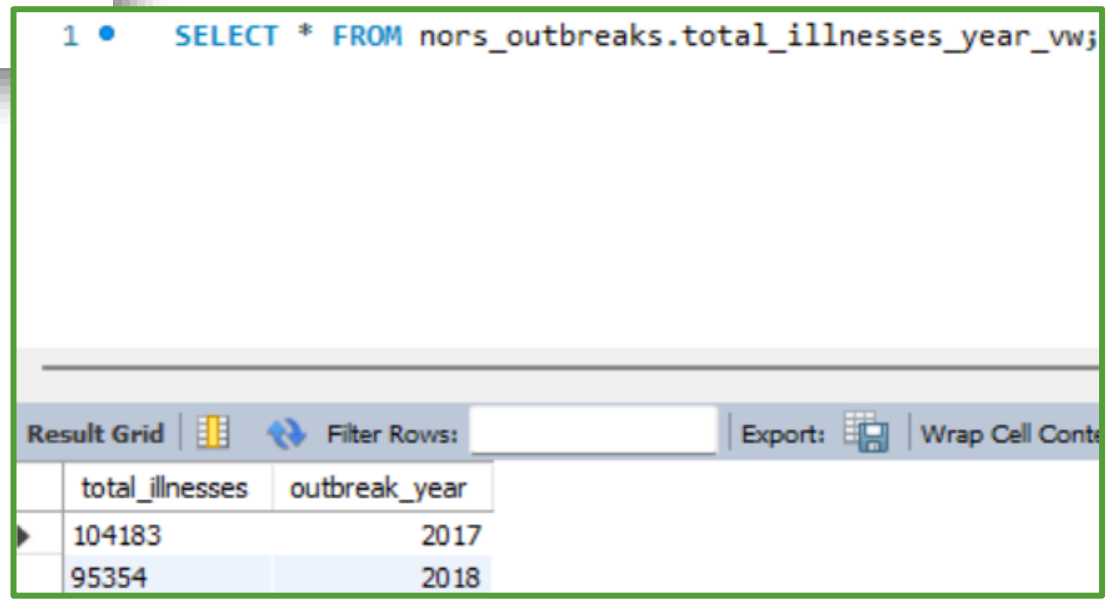
Analytical Research Question	Expected Results
1. How many total illnesses were reported to the National Outbreak Response System (NORS) in 2017 compared to 2018, grouped by state and outbreak year? Include outbreak year and total number of illnesses. Report the answer as a view.	A static table containing number of illnesses reported in 2017 and in 2018
2. Which pathogens cause the most and least number of illnesses in 2017? Create a list of pathogens and total illnesses in descending order of total illnesses. Conduct the same analysis for 2018.	Two CSV exports of the list of total pathogens in 2017 and 2018.
3. What transmission mode led to the most outbreaks in 2017? What about in 2018?	Two CSVs containing the number of outbreaks, transmission mode and year
4. Create a flag for all norovirus related outbreaks from 2017-2018. How many norovirus outbreaks occurred in California that led to more than 10 illnesses?	A new table and two CSVs documenting the answers for both years
5. How many unique food vehicles were associated with more hospitalizations than outbreak_id 982 in 2017? What about in 2018?	The count of distinct food vehicles in 2017 and 2018

Question 1: How many total illnesses were reported to the National Outbreak Response System (NORS) in 2017 compared to 2018, grouped by state and outbreak year? Include outbreak year and total number of illnesses. Report the answer as a view.

Query:

```
CREATE VIEW total_illnesses_year_vw AS
SELECT SUM(num_illnesses) AS total_illnesses, outbreak_year
FROM timeframe
LEFT JOIN outcomes
ON timeframe.outbreak_id = outcomes.outbreak_id
GROUP BY outbreak_year;
```

Results: 2017 had 104,183 illnesses. 2018 had 95,354 illnesses



Question 2: Which pathogens cause the most and least number of illnesses in 2017? Create a list of pathogens and total illnesses in descending order of total illnesses. Conduct the same analysis for 2018.

Queries for both 2017 and 2018:

```sql
-- 2017;
SELECT SUM(num_illnesses) as total_illnesses, pathogen
FROM timeframe
INNER JOIN etiology
ON timeframe.outbreak_id = etiology.outbreak_id
INNER JOIN outcomes
ON timeframe.outbreak_id = outcomes.outbreak_id
WHERE outbreak_year = 2017
GROUP BY pathogen
ORDER BY total_illnesses DESC;
```

```sql
-- 2018;
SELECT SUM(num_illnesses) as total_illnesses, pathogen
FROM timeframe
INNER JOIN etiology
ON timeframe.outbreak_id = etiology.outbreak_id
INNER JOIN outcomes
ON timeframe.outbreak_id = outcomes.outbreak_id
WHERE outbreak_year = 2018
GROUP BY pathogen
ORDER BY total_illnesses DESC;
```

Results: Norovirus caused the most illnesses in 2017, while Norovirus unknown caused the most illnesses in 2018.

total_illnesses	pathogen
23118	Norovirus
22417	Norovirus Genogroup II
18039	Norovirus unknown
14488	
4921	Norovirus Genogroup I
4219	Salmonella enterica
3166	Norovirus Genogroup II; Norovirus Genogroup II

total_illnesses	pathogen
21419	Norovirus unknown
19224	Norovirus Genogroup II
17752	Norovirus
10522	
5071	Salmonella enterica
3954	Norovirus Genogroup I
2512	Norovirus Genogroup II; Norovirus Genogroup II

Question 3: What transmission mode led to the most outbreaks in 2017? What about in 2018?

Queries for both 2017 and 2018:

```
-- 2017
SELECT COUNT(transmission_mode) as num_outbreaks, transmission_mode, outbreak_year
FROM etiology
LEFT JOIN timeframe
ON etiology.outbreak_id = timeframe.outbreak_id
WHERE outbreak_year = 2017
GROUP BY transmission_mode
ORDER BY num_outbreaks DESC;
```

```
-- 2018
SELECT COUNT(transmission_mode) as num_outbreaks, transmission_mode, outbreak_year
FROM etiology
LEFT JOIN timeframe
ON etiology.outbreak_id = timeframe.outbreak_id
WHERE outbreak_year = 2018
GROUP BY transmission_mode
ORDER BY num_outbreaks DESC;
```

Results: Person-to-person transmission mode led to the greatest number of outbreaks for both 2017 and 2018.

num_outbreaks	transmission_mode	outbreak_year
2551	Person-to-person	2017
861	Food	2017
530	Indeterminate/Other/Unknown	2017
112	Water	2017
61	Animal Contact	2017
16	Environmental contamination other than food/w...	2017

num_outbreaks	transmission_mode	outbreak_year
2397	Person-to-person	2018
1054	Food	2018
488	Indeterminate/Other/Unknown	2018
138	Water	2018
61	Animal Contact	2018
7	Environmental contamination other than food/w...	2018

Question 4: Create a flag for all norovirus related outbreaks from 2017-2018. How many norovirus outbreaks occurred in California that led to more than 10 illnesses?

Query:

```
CREATE TABLE etiology_flag
SELECT etiology_id, transmission_mode, serotype, pathogen, etiology_status, outbreak_id,
    CASE
        WHEN pathogen LIKE "%Noro%" THEN "Yes"
        ELSE "No"
    END AS norovirus_flag
FROM etiology;

SELECT e.outbreak_id, t.outbreak_year, t.state, o.num_illnesses, e.pathogen
FROM etiology_flag AS e, outcomes AS o, timeframe AS t, situation AS s
WHERE t.outbreak_id = o.outbreak_id AND
    t.outbreak_id = e.outbreak_id AND
    t.outbreak_id = s.outbreak_id AND
    t.state LIKE "%Calif%" AND
    o.num_illnesses > 10 AND
    e.norovirus_flag = "Yes";
```

Results: New table with norovirus_flag and the following data that met the query criterion

etiology_id	transmission_mode	serotype	pathogen	etiology_status	outbreak_id	norovirus_flag
1	Person-to-person		Norovirus	Suspected	1	Yes
2	Person-to-person	GII.4 Sydney	Norovirus Genogroup II	Confirmed	2	Yes
3	Person-to-person		Norovirus	Suspected	3	Yes
4	Person-to-person		Norovirus	Suspected	4	Yes
5	Person-to-person		Rotavirus; Sapovirus	Confirmed; Suspected	5	No
6	Person-to-person		Norovirus	Suspected	6	Yes
7	Person-to-person	GII.4 Sydn...	Norovirus Genogroup ...	Confirmed; Suspected	7	Yes
8	Person-to-person		Norovirus Genogroup II	Confirmed	8	Yes
9	Person-to-person		Norovirus	Suspected	9	Yes
10	Person-to-person	GII.6;	Norovirus Genogroup ...	Suspected; Suspected	10	Yes

outbreak_id	outbreak_year	state	num_illnesses	pathogen
1750	2017	California	33	Norovirus Genogroup II
2863	2017	California	31	Norovirus Genogroup ...
2864	2017	California	29	Norovirus unknown
3344	2017	California	27	Norovirus unknown
3351	2017	California	15	Norovirus unknown
4300	2017	California	13	Norovirus unknown
4330	2017	California	19	Norovirus Genogroup II
4359	2017	California	79	Norovirus unknown
4653	2017	California	40	Norovirus unknown

Question 5: How many unique food vehicles were associated with more hospitalizations than outbreak_id 982 in 2017? What about in 2018? Create both as a view.

Queries for both 2017 and 2018:

```
-- 2017;
CREATE VIEW 2017_food_vehicles_vw AS
(SELECT COUNT(DISTINCT(food_vehicle)) AS num_distinct_foods
FROM situation, outcomes, timeframe
WHERE situation.outbreak_id = outcomes.outbreak_id AND
    situation.outbreak_id = timeframe.outbreak_id AND
    num_hospitalizations > (SELECT num_hospitalizations
                            FROM outcomes
                            WHERE outbreak_id = 982) AND
    outbreak_year = 2017);
```

```
-- 2018;
CREATE VIEW 2018_food_vehicles_vw AS
SELECT COUNT(DISTINCT(food_vehicle)) AS num_distinct_foods
FROM situation, outcomes, timeframe
WHERE situation.outbreak_id = outcomes.outbreak_id AND
    situation.outbreak_id = timeframe.outbreak_id AND
    num_hospitalizations > (SELECT num_hospitalizations
                            FROM outcomes
                            WHERE outbreak_id = 982) AND
    outbreak_year = 2018;
```

Results: In 2017, 12 distinct foods met the criteria while 19 met the criteria in 2018.

num_distinct_foods
12

num_distinct_foods
19

MySQL Commands

- CREATE SCHEMA
- CREATE TABLE
- PRIMARY KEY
- INSERT INTO
- SELECT FROM
- INNER JOIN
- GROUP BY
- WHERE
- COUNT
- SUM
- ORDER BY
- CASE WHEN
- CREATE VIEW
- DESC
- DISTINCT

Discussion & Conclusions

- General decrease in the frequency and intensity of foodborne disease outbreaks in 2018 compared to 2017
  - Number of illnesses
  - Number of outbreaks
  - Number of illnesses associated with norovirus
- Future directions:
  - Expand timeframe in which data is compared to allow local jurisdictions time to implement regulations
  - Gather more detailed data on transmission mode to determine any relationship with the 2017 Food Code new section
  - Perform statistical testing on the quantitative measures included in NORS to determine statistical differences

References

- FDA Food Code. FDA. Published February 8, 2023. Accessed April 19, 2023. https://www.fda.gov/food/retail-food-protection/fda-food-code
- New 2017 Food Code Section on Bandages, Finger Cots, or Finger Stalls. Accessed April 22, 2023. https://www.fda.gov/media/111463/download
- National Outbreak Reporting System (NORS) Dashboard | CDC. Accessed April 20, 2023. https://wwwn.cdc.gov/norsdashboard/