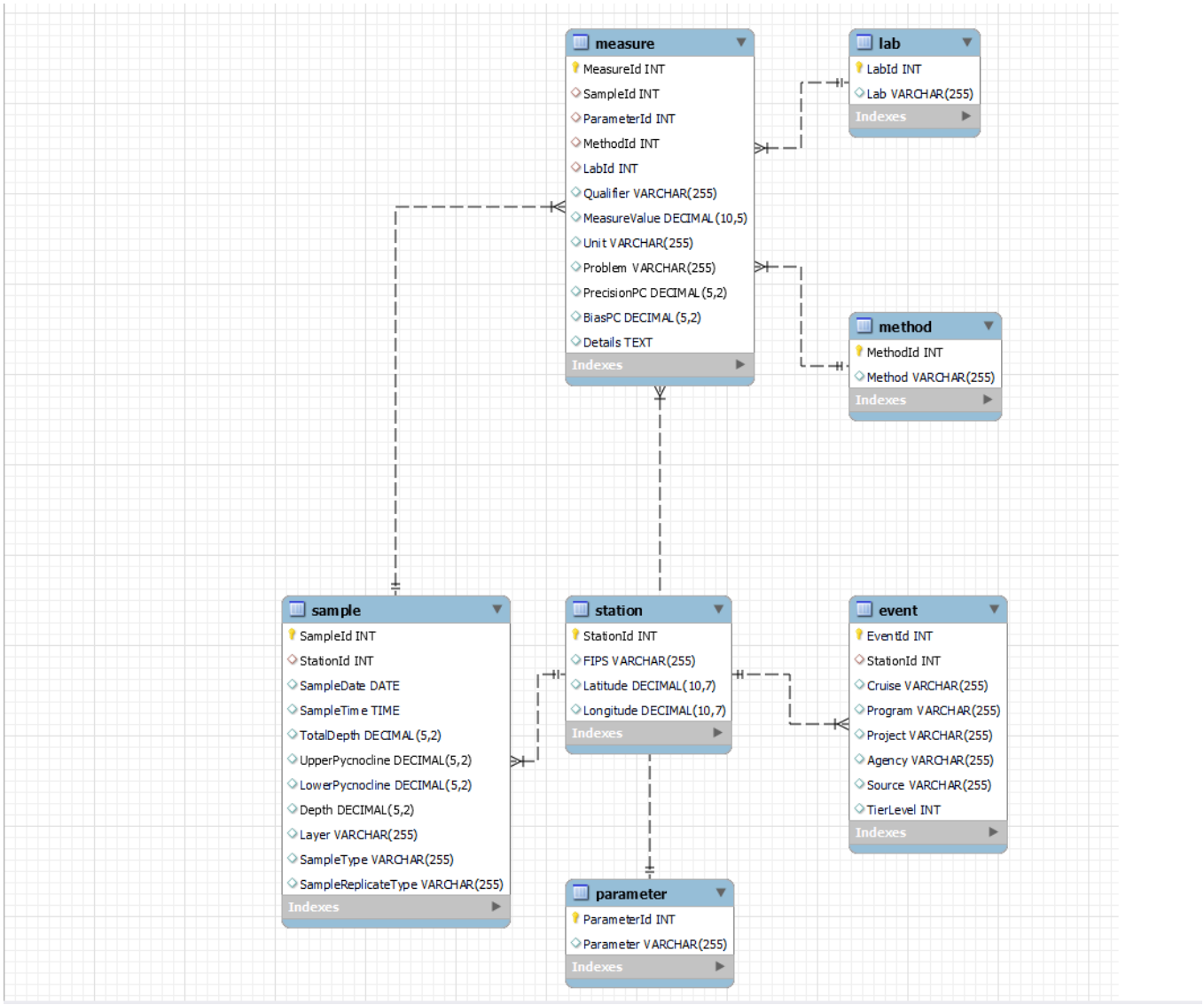
Course Project cs450
Due: Wed Apr 24, 2024 11:59pm

```
CREATE TABLE `station` (
  `StationId` INT NOT NULL AUTO_INCREMENT,
  `FIPS` VARCHAR(255),
  `Latitude` DECIMAL(10, 8),
  `Longitude` DECIMAL(11, 8),
  PRIMARY KEY (`StationId`)
);

CREATE TABLE `event` (
  `EventId` INT NOT NULL AUTO_INCREMENT,
  `StationId` INT,
  `Cruise` VARCHAR(255),
  `Program` VARCHAR(255),
  `Project` VARCHAR(255),
  `Agency` VARCHAR(255),
  `Source` VARCHAR(255),
  `TierLevel` VARCHAR(255),
  PRIMARY KEY (`EventId`),
  FOREIGN KEY (`StationId`) REFERENCES `station`(`StationId`)
);

CREATE TABLE `parameter` (
  `ParameterId` INT NOT NULL AUTO_INCREMENT,
  `Parameter` VARCHAR(255),
  PRIMARY KEY (`ParameterId`)
);


CREATE TABLE `method` (
  `MethodId` INT AUTO_INCREMENT PRIMARY KEY,
  `Method` VARCHAR(255)
);

CREATE TABLE `lab` (
  `LabId` INT NOT NULL AUTO_INCREMENT,
  `Lab` VARCHAR(255),
  PRIMARY KEY (`LabId`)
);
```

```
CREATE TABLE sample (
SampleId INT NOT NULL AUTO_INCREMENT,
StationId INT,
SampleDate DATE,
SampleTime TIME,
TotalDepth DECIMAL(5, 2),
UpperPycnocline DECIMAL(5, 2),
LowerPycnocline DECIMAL(5, 2),
Depth DECIMAL(5, 2),
Layer VARCHAR(255),
SampleType VARCHAR(255),
SampleReplicateType VARCHAR(255),
PRIMARY KEY (SampleId),
FOREIGN KEY (StationId) REFERENCES station (StationId)
);


CREATE TABLE `measure` (
  `MeasureId` INT NOT NULL AUTO_INCREMENT,
  `SampleId` INT,
  `ParameterId` INT,
  `MethodId` INT,
  `LabId` INT,
  `Qualifier` VARCHAR(255),
  `MeasureValue` DECIMAL(10, 4),
  `Unit` VARCHAR(255),
  `Problem` VARCHAR(255),
  `PrecisionPC` DECIMAL(5, 2),
  `BiasPC` DECIMAL(5, 2),
  `Details` VARCHAR(255),
  PRIMARY KEY (`MeasureId`),
  FOREIGN KEY (`SampleId`) REFERENCES `sample`(`SampleId`),
  FOREIGN KEY (`ParameterId`) REFERENCES `parameter`(`ParameterId`),
  FOREIGN KEY (`MethodId`) REFERENCES `method`(`MethodId`),
  FOREIGN KEY (`LabId`) REFERENCES `lab`(`LabId`)
);
```

2)


measure
MeasureId INT
SampleId INT
ParameterId INT
MethodId INT
LabId INT
Qualifier VARCHAR(255)
MeasureValue DECIMAL(10,5)
Unit VARCHAR(255)
Problem VARCHAR(255)
PrecisionPC DECIMAL(5,2)
BiasPC DECIMAL(5,2)
Details TEXT
Indexes
lab
LabId INT
Lab VARCHAR(255)
Indexes
method
MethodId INT
Method VARCHAR(255)
Indexes
sample
SampleId INT
StationId INT
SampleDate DATE
SampleTime TIME
TotalDepth DECIMAL(5,2)
UpperPycnocline DECIMAL(5,2)
LowerPycnocline DECIMAL(5,2)
Depth DECIMAL(5,2)
Layer VARCHAR(255)
SampleType VARCHAR(255)
SampleReplicateType VARCHAR(255)
Indexes
station
StationId INT
FIPS VARCHAR(255)
Latitude DECIMAL(10,7)
Longitude DECIMAL(10,7)
Indexes
event
EventId INT
StationId INT
Cruise VARCHAR(255)
Program VARCHAR(255)
Project VARCHAR(255)
Agency VARCHAR(255)
Source VARCHAR(255)
TierLevel INT
Indexes
parameter
ParameterId INT
Parameter VARCHAR(255)
Indexes

3)

Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap C

EventId	StationId	Cruise	Program	Project	Agency	Source	TierLevel
1	1	NTN021	NTWQM	NTN	DEDNREC	DEDNREC	T3
2	1	NTN021	NTWQM	NTN	DEDNREC	DEDNREC	T3
3	1	NTN021	NTWQM	NTN	DEDNREC	DEDNREC	T3
4	1	NTN021	NTWQM	NTN	DEDNREC	DEDNREC	T3
5	1	NTN021	NTWQM	NTN	DEDNREC	DEDNREC	T3
NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL

SampleId	StationId	SampleDate	SampleTime	TotalDepth	UpperPycnocline	LowerPycnocline	Depth	Layer	SampleType	SampleReplicateType
1	1	2021-01-28	10:36:00	NULL	NULL	NULL	0.00	S	HVIC	S1
2	1	2021-02-02	10:15:00	NULL	NULL	NULL	0.00	S	HVIC	S1
3	1	2021-02-24	09:17:00	NULL	NULL	NULL	0.00	S	HVIC	S1
4	1	2021-03-10	09:03:00	NULL	NULL	NULL	0.00	S	HVIC	S1
5	1	2021-04-20	08:41:00	NULL	NULL	NULL	0.00	S	HVIC	S1
NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL

Result Grid | Filter Rows:

ParameterId	Parameter
1	CHLA
2	CLW
3	DO
4	DO_SAT_P
5	DOC
NULL	NULL

4)

Q1: What is the average CHLA (chlorophyll a) concentration for each month? How does it compare to the overall average CHLA value for the entire dataset? (Hint: Use the column MeasureValue for CHLA concentration.)

```
SELECT
  MONTH(s.SampleDate) AS Month,
  AVG(m.MeasureValue) AS AvgCHLA,
  (SELECT AVG(MeasureValue)
   FROM Measure
   WHERE ParameterId IN (SELECT ParameterId FROM Parameter WHERE Parameter =
'CHLA')) AS OverallAvgCHLA
FROM Measure m
JOIN Sample s ON m.SampleId = s.SampleId
JOIN Parameter p ON m.ParameterId = p.ParameterId
WHERE p.Parameter = 'CHLA'
GROUP BY MONTH(s.SampleDate);
```

Q2 What is the maximum and minimum CHLA concentration at each station, and when (date and time) were the measurements taken? (Hint: Use the column MeasureValue for CHLA concentration.)

```
SELECT
  s.StationId,
  MAX(m.MeasureValue) AS MaxCHLA,
  MIN(m.MeasureValue) AS MinCHLA,
  (SELECT SampleDate FROM Sample WHERE SampleId = (SELECT SampleId FROM
Measure WHERE MeasureValue = MAX(m.MeasureValue) LIMIT 1)) AS MaxCHLADate,
  (SELECT SampleTime FROM Sample WHERE SampleId = (SELECT SampleId FROM
Measure WHERE MeasureValue = MAX(m.MeasureValue) LIMIT 1)) AS MaxCHLATime,
  (SELECT SampleDate FROM Sample WHERE SampleId = (SELECT SampleId FROM
Measure WHERE MeasureValue = MIN(m.MeasureValue) LIMIT 1)) AS MinCHLADate,
  (SELECT SampleTime FROM Sample WHERE SampleId = (SELECT SampleId FROM
Measure WHERE MeasureValue = MIN(m.MeasureValue) LIMIT 1)) AS MinCHLATime
FROM Measure m
JOIN Sample sa ON m.SampleId = sa.SampleId
JOIN Station s ON sa.StationId = s.StationId
JOIN Parameter p ON m.ParameterId = p.ParameterId
WHERE p.Parameter = 'CHLA'
GROUP BY s.StationId;
```

Q3: How many samples were taken for each SampleReplicateType (S1, FS1, etc.) across all stations?

```
SELECT
  SampleReplicateType,
  COUNT(*) AS SampleCount
FROM Sample
GROUP BY SampleReplicateType;
```

Q4: Are there any stations where the CHLA value consistently does not exceed a certain threshold (18.0 ug/L) throughout the year? If so, which stations, and during which months does this occur?

```
SELECT
  s.StationId,
  GROUP_CONCAT(DISTINCT MONTH(sa.SampleDate) ORDER BY
MONTH(sa.SampleDate)) AS Months
FROM Station s
JOIN Sample sa ON s.StationId = sa.StationId
JOIN Measure m ON sa.SampleId = m.SampleId
JOIN Parameter p ON m.ParameterId = p.ParameterId
WHERE p.Parameter = 'CHLA' AND m.MeasureValue <= 18.0
GROUP BY s.StationId
HAVING COUNT(DISTINCT MONTH(sa.SampleDate)) = 12;
```