PROJECT REPORT 1st MILESTONE

Manik Jain
(2022MCS2832)

Sagar Agrawal
(2022MCS2065)

Vinit Chandak
(2022EET2109)

20/03/2023

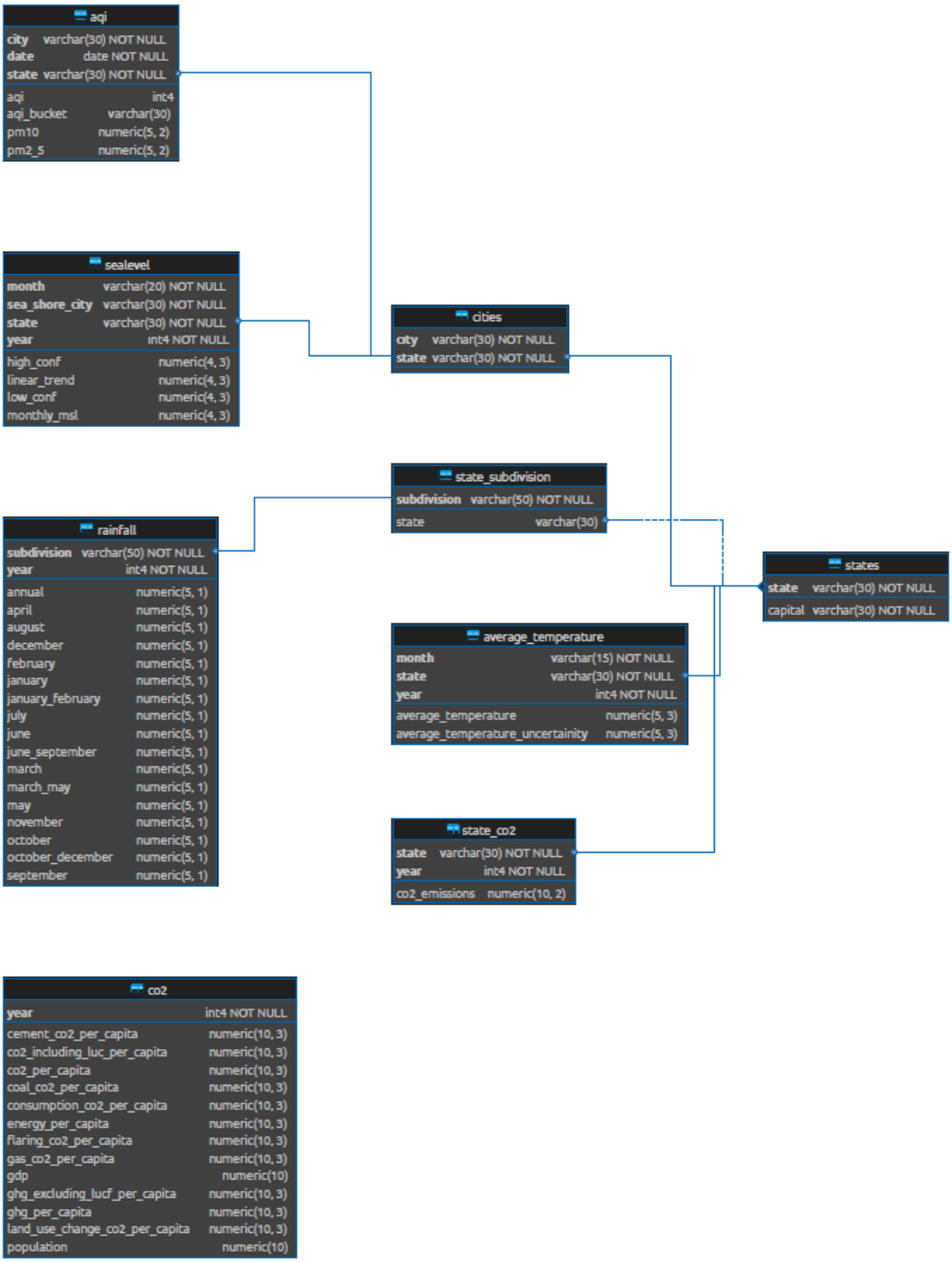
—

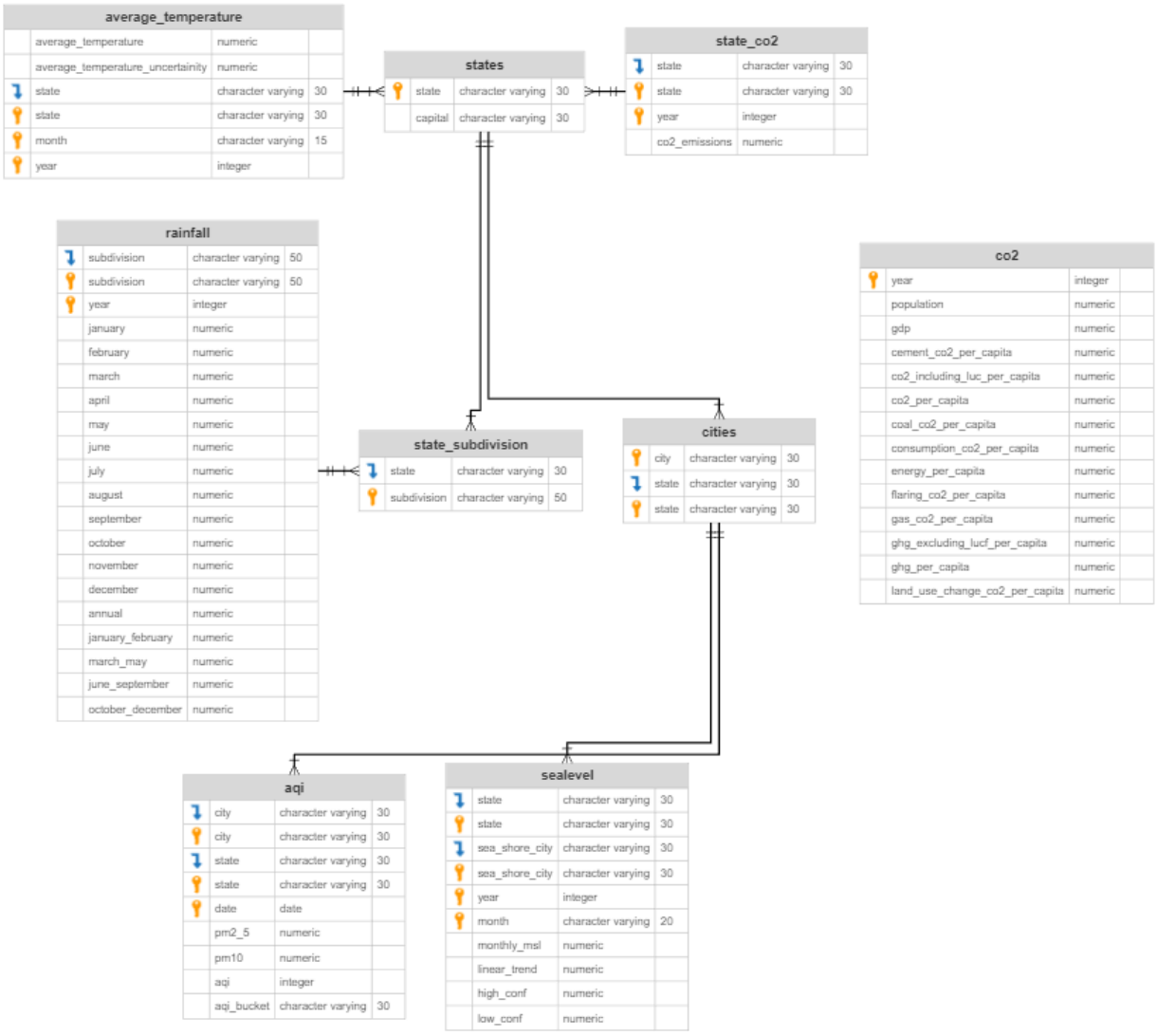
Introduction to Database Systems
(COL 362/632)

—

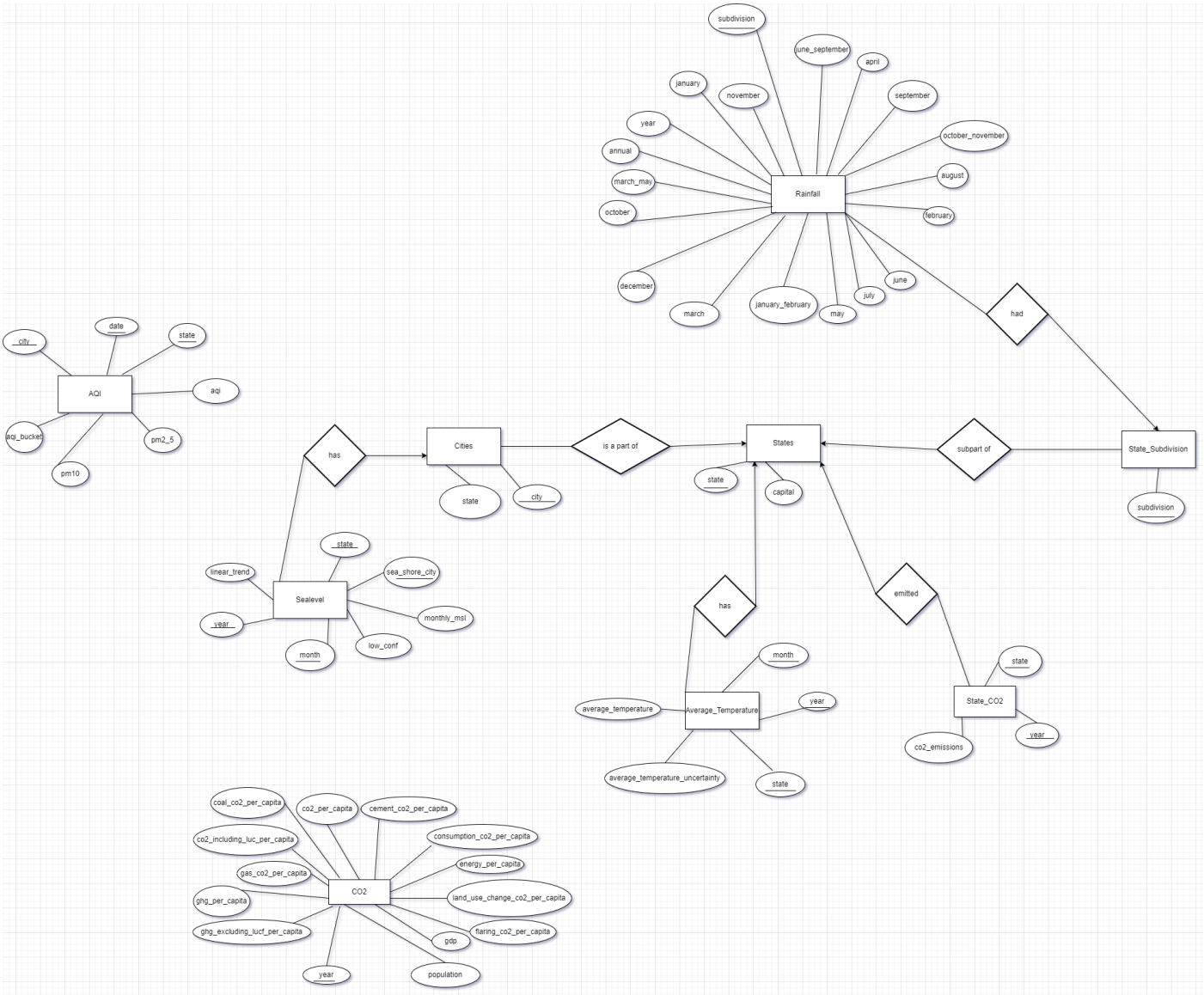
Prof. Srikanta Bedathur

ENTITY RELATIONSHIP DIAGRAMS:




average_temperature
average_temperature numeric
average_temperature_uncertainity numeric
state character varying 30
state character varying 30
month character varying 15
year integer
states
state character varying 30
capital character varying 30
state_co2
state character varying 30
state character varying 30
year integer
co2_emissions numeric
rainfall
subdivision character varying 50
subdivision character varying 50
year integer
january numeric
february numeric
march numeric
april numeric
may numeric
june numeric
july numeric
august numeric
september numeric
october numeric
november numeric
december numeric
annual numeric
january_february numeric
march_may numeric
june_september numeric
october_december numeric
state_subdivision
state character varying 30
subdivision character varying 50
cities
city character varying 30
state character varying 30
state character varying 30
co2
year integer
population numeric
gdp numeric
cement_co2_per_capita numeric
co2_including_luc_per_capita numeric
co2_per_capita numeric
coal_co2_per_capita numeric
consumption_co2_per_capita numeric
energy_per_capita numeric
flaring_co2_per_capita numeric
gas_co2_per_capita numeric
ghg_excluding_lucf_per_capita numeric
ghg_per_capita numeric
land_use_change_co2_per_capita numeric
aqi
city character varying 30
city character varying 30
state character varying 30
state character varying 30
date date
pm2_5 numeric
pm10 numeric
aqi integer
aqi_bucket character varying 30
sealevel
state character varying 30
state character varying 30
sea_shore_city character varying 30
sea_shore_city character varying 30
year integer
month character varying 20
monthly_msl numeric
linear_trend numeric
high_conf numeric
low_conf numeric

ER DIAGRAM:


subdivision
june_september
april
january
november
september
year
annual
october_november
march_may
august
Rainfall
october
february
december
march
january_february
may
july
june
had
city
date
state
AQI
aqi
aqi_bucket
pm2_5
pm10
has
Cities
is a part of
States
subpart of
State_Subdivision
state
city
state
capital
subdivision
state
linear_trend
sea_shore_city
Sealevel
year
monthly_msl
month
low_conf
has
emitted
month
state
average_temperature
Average_Temperature
year
State_CO2
co2_emissions
year
average_temperature_uncertainty
state
coal_co2_per_capita
co2_per_capita
cement_co2_per_capita
co2_including_luc_per_capita
consumption_co2_per_capita
energy_per_capita
gas_co2_per_capita
CO2
ghg_per_capita
land_use_change_co2_per_capita
ghg_excluding_luc_per_capita
flaring_co2_per_capita
gdp
year
population

FUNCTIONAL DEPLENDENCIES:

state_capital:

state → capital

state_subdivision:

subdivision → state

average_temperature:

{month, state, year} → average_temperature, average_temperature_uncertainity

state_co2:

{state, year} → co2_emissions

aqi:

{city, date, state} → aqi, aqi_bucket, pm10, pm2_5

sealevel:

{month, sea_shore_city, state, year} → high_conf, linear_trend, low_conf, monthly_msl

rainfall:

{subdivision, year} → january,
february,
march,
april,
may,
june,
july,
august,

september,
october,
november,
december,
annual,
january_february,
march_may,
june_september,
october_december

co2:

year → population,
gdp,
cement_co2_per_capita ,
co2_including_luc_per_capita ,
co2_per_capita ,
coal_co2_per_capita ,
consumption_co2_per_capita ,
energy_per_capita ,
flaring_co2_per_capita ,
gas_co2_per_capita ,
ghg_excluding_lucf_per_capita ,
ghg_per_capita,
land_use_change_co2_per_capita

NORMALIZATION:

1. Rainfall

Before Normalization:

{subdivision, year, state} → january, february, march, april, may, june, july, august, september, october, november, december, annual, january_february, march_may, june_september, october_december

subdivision → state

The given relation is in 3NF.
Because of the FD subdivision → state, the relation is not in BCNF as subdivision is not a superkey.

After Normalization:

Rainfall:
{subdivision, year} → january, february, march, april, may, june, july, august, september, october, november, december, annual, january_february, march_may, june_september, october_december

state_subdivision:
subdivision → state

Both the new relations are in BCNF as there is only one FD where LHS is a superkey.

All of the other relations were already in proper normal forms and hence there was no need to normalize them. **There are not a lot of dependencies in our database because of the inherent nature of the dataset**.

GITHUB REPOSITORY:

https://github.com/Vinit-Chandak/SMV