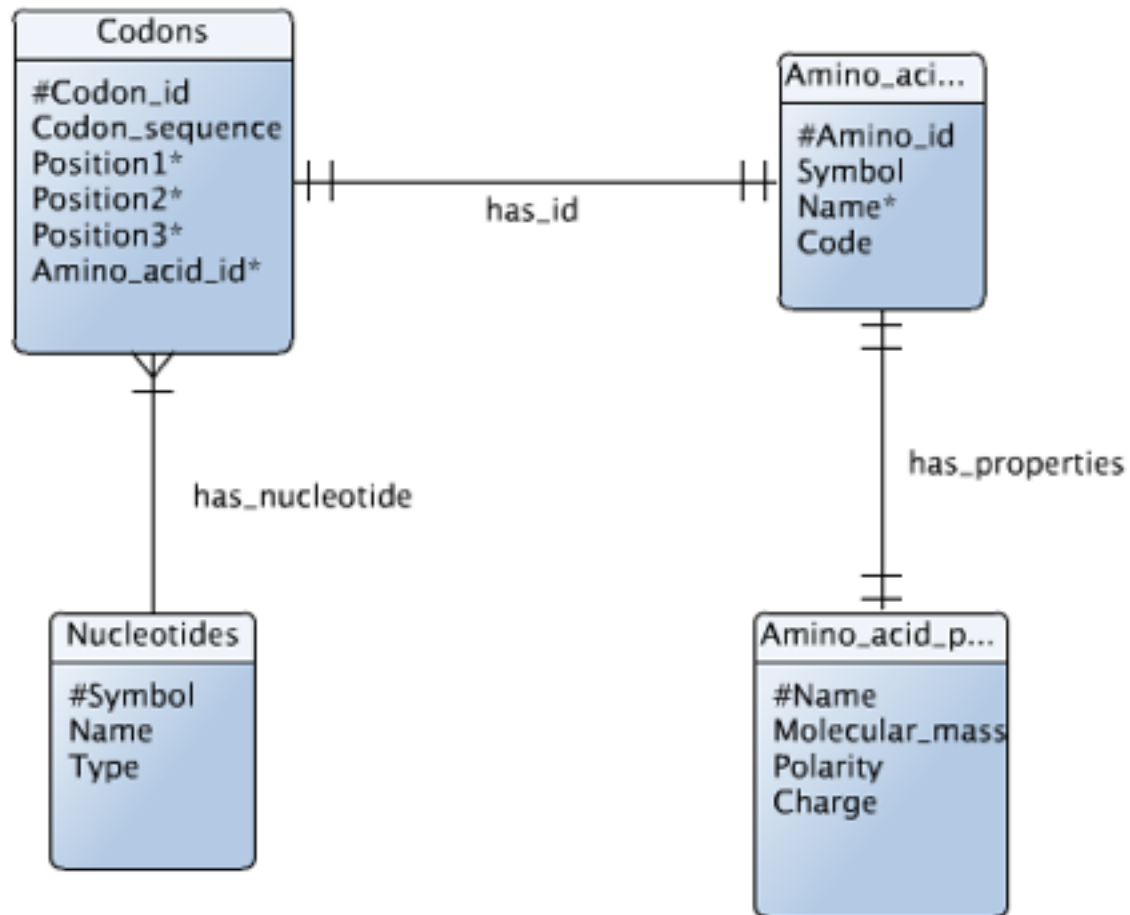1)

Some assumptions: From the slides it seems that in ER the convention for this course is just to name the relationships between entities instead of using the "relationship"-box and two links.
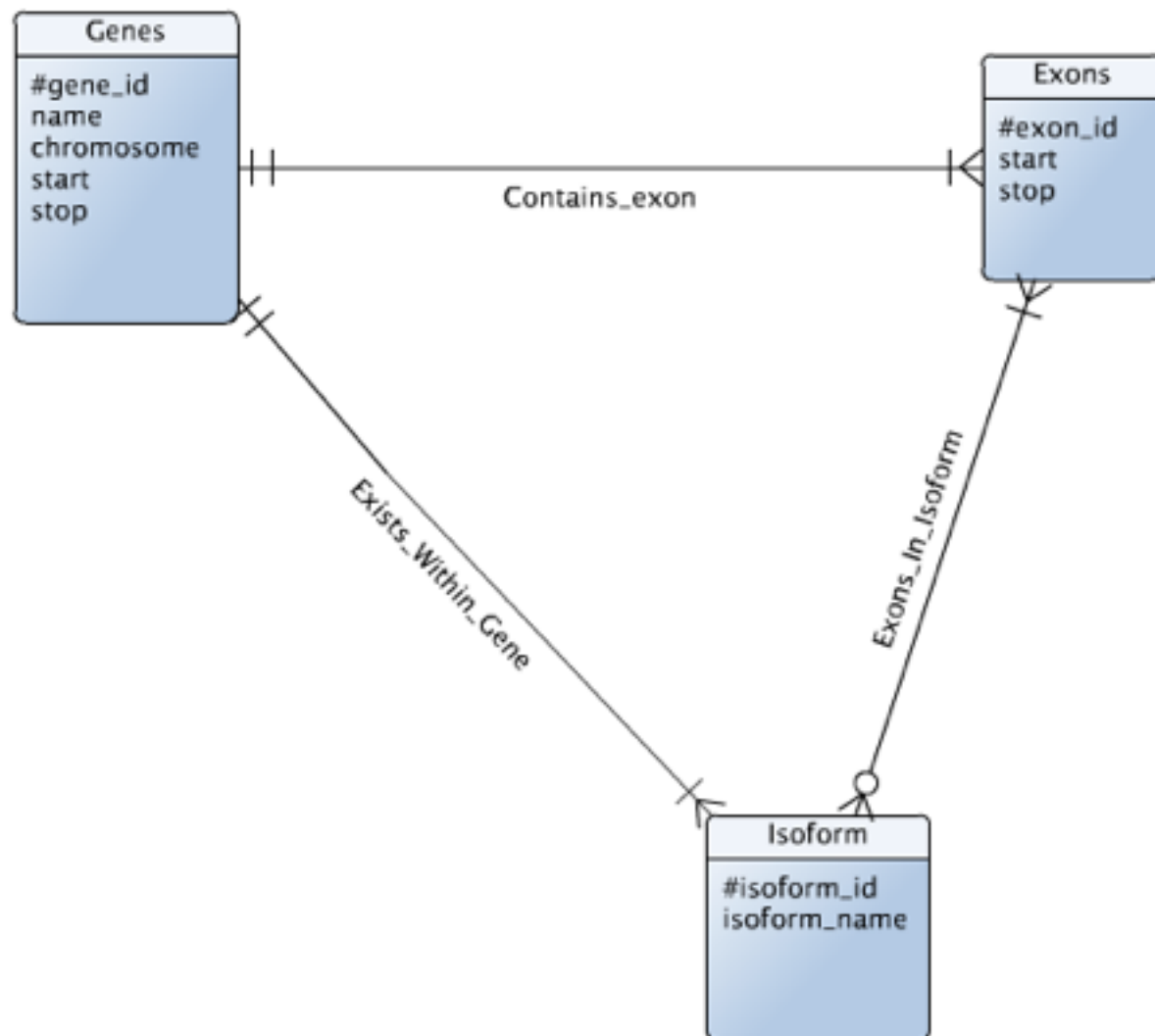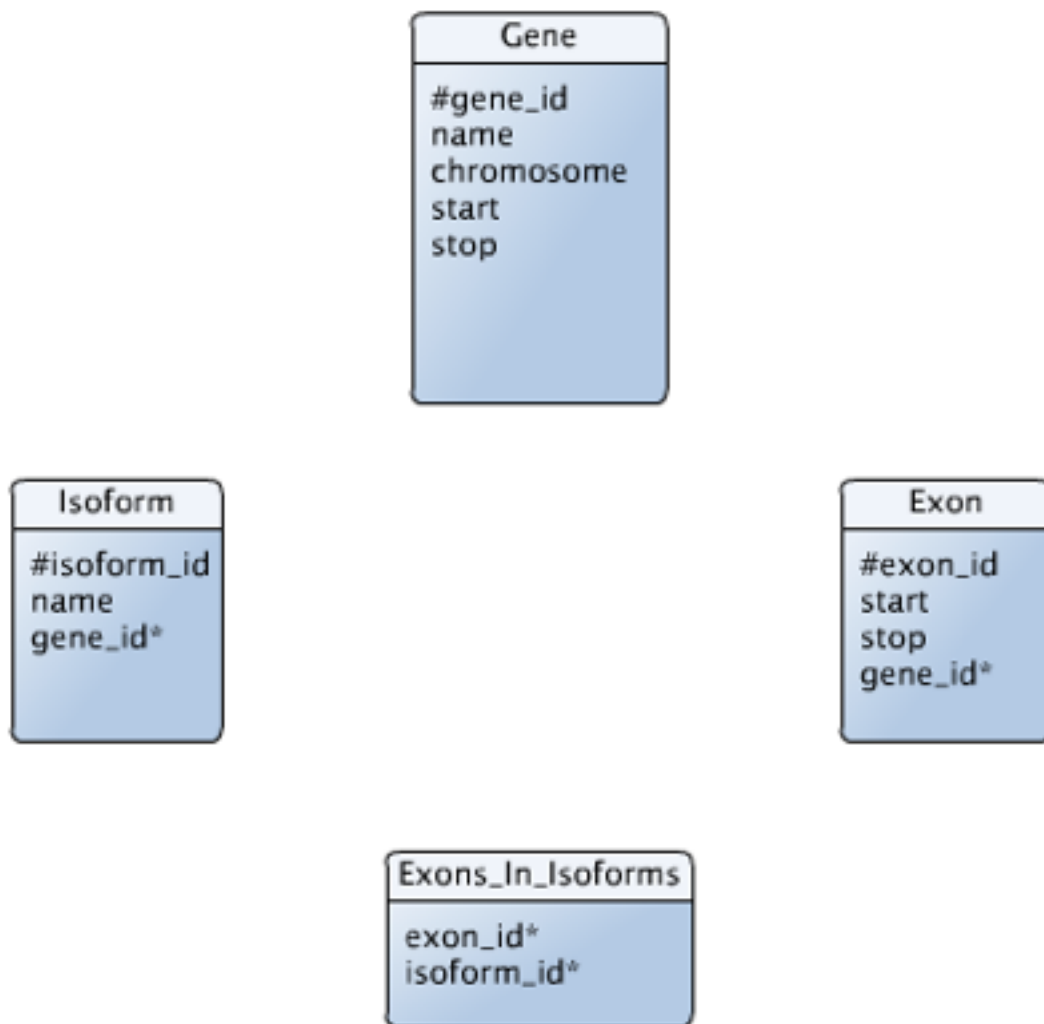
2i)
Genes, Exons and Isoforms (disregarding the two coupling entities needed to make the
relationships work)

2ii)
Disregarding coupling entity between exons and isoform for this

Genes

#gene_id
name
chromosome
start
stop

Exons

#exon_id
start
stop

Contains_exon

Exists_Within_Gene

Exons_In_Isoform

Isoform

#isoform_id
isoform_name

2iii)

Gene

#gene_id
name
chromosome
start
stop

Isoform

#isoform_id
name
gene_id*

Exon

#exon_id
start
stop
gene_id*

Exons_In_Isoforms

exon_id*
isoform_id*

Exons_In_Isoforms is a weak entity keeping track of all the exons within each isoform.

3i)
Gene, Chromosome, Publication, Author

Assuming a journal just has PK as name so there is no point making a separate table for it.

3ii)
Cardinality assumptions:
*Each gene exists one one and only one chromosome
*Each chromosome has at least one gene
*Each gene is referenced by at least one publication
*Each publication references at least one gene
*Each publication has at least one author
*An author can be the author of zero to many publications

Non-atomic values, references*, authors* is due to it not being normalized yet.
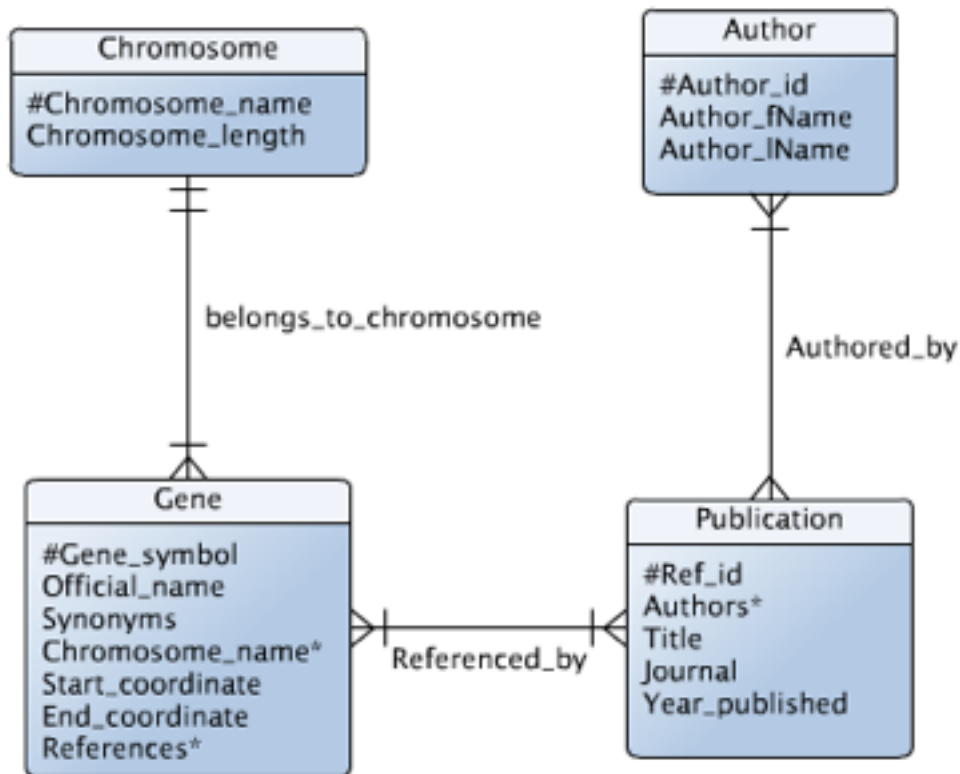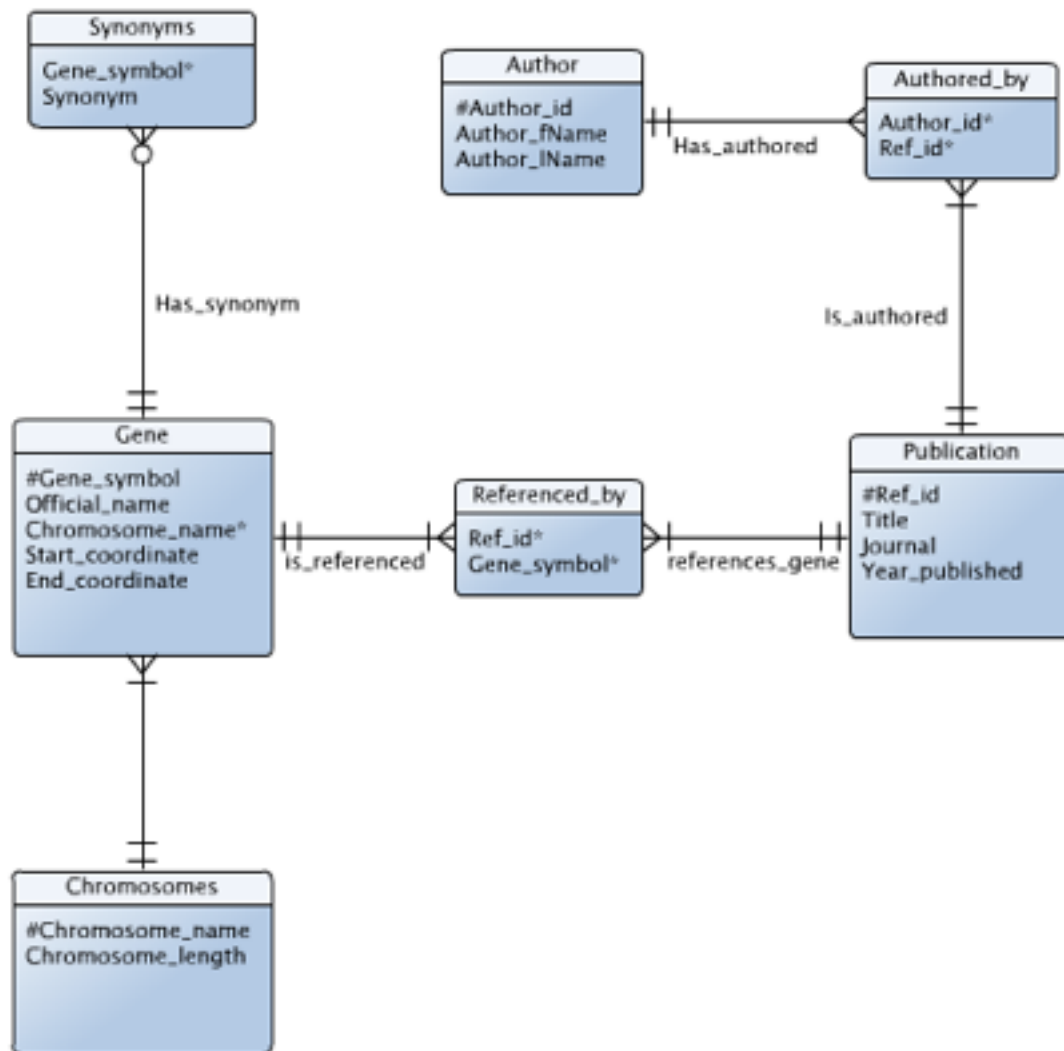
3iii)

Only fields to be dealt with is: References*, Authors* and Synonyms
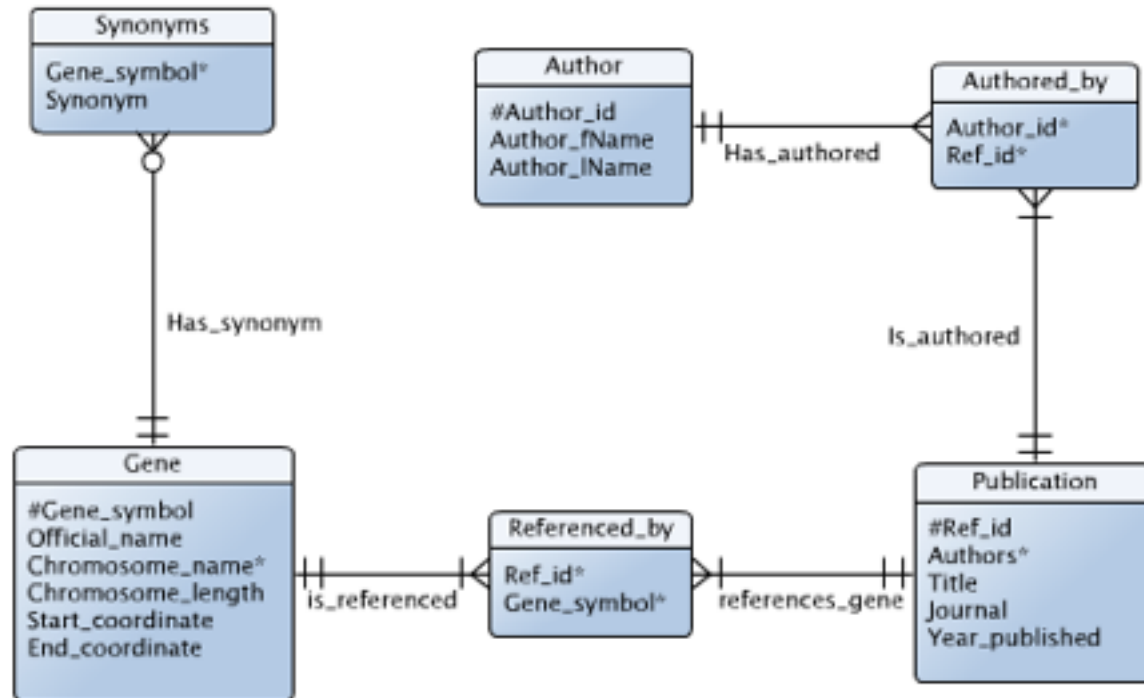
This is 1NF because it has no multi-value attributes anywhere, plus it is not 2NF because Chromosome_length depends on Chromosome_name.

3iv)
Not sure if this is right because of Chromosome_name depending on Gene_symbol. Would like
feedback/clarification on that if possible!

Synonyms
Gene_symbol*
Synonym

Author
#Author_id
Author_fName
Author_IName

Has_authored

Authored_by
Author_id*
Ref_id*

Has_synonym

Is_authored

Gene
#Gene_symbol
Official_name
Chromosome_name*
Start_coordinate
End_coordinate

is_referenced

Referenced_by
Ref_id*
Gene_symbol*

references_gene

Publication
#Ref_id
Title
Journal
Year_published

Chromosomes
#Chromosome_name
Chromosome_length

4)

Cardinality assumptions: Each Cruise has only one ship. A ship doesn't necessarily need to be in use, or it can be scheduled for use for multiple cruises. A cruise has one start port, and one or more ports in its route.
Other notes: The relation has_route doesn't include arrival_time, departure_time, but if this wasn't ER Route would be a weak entity with the attributes cruise_id, arrival_time, departure_time, port.
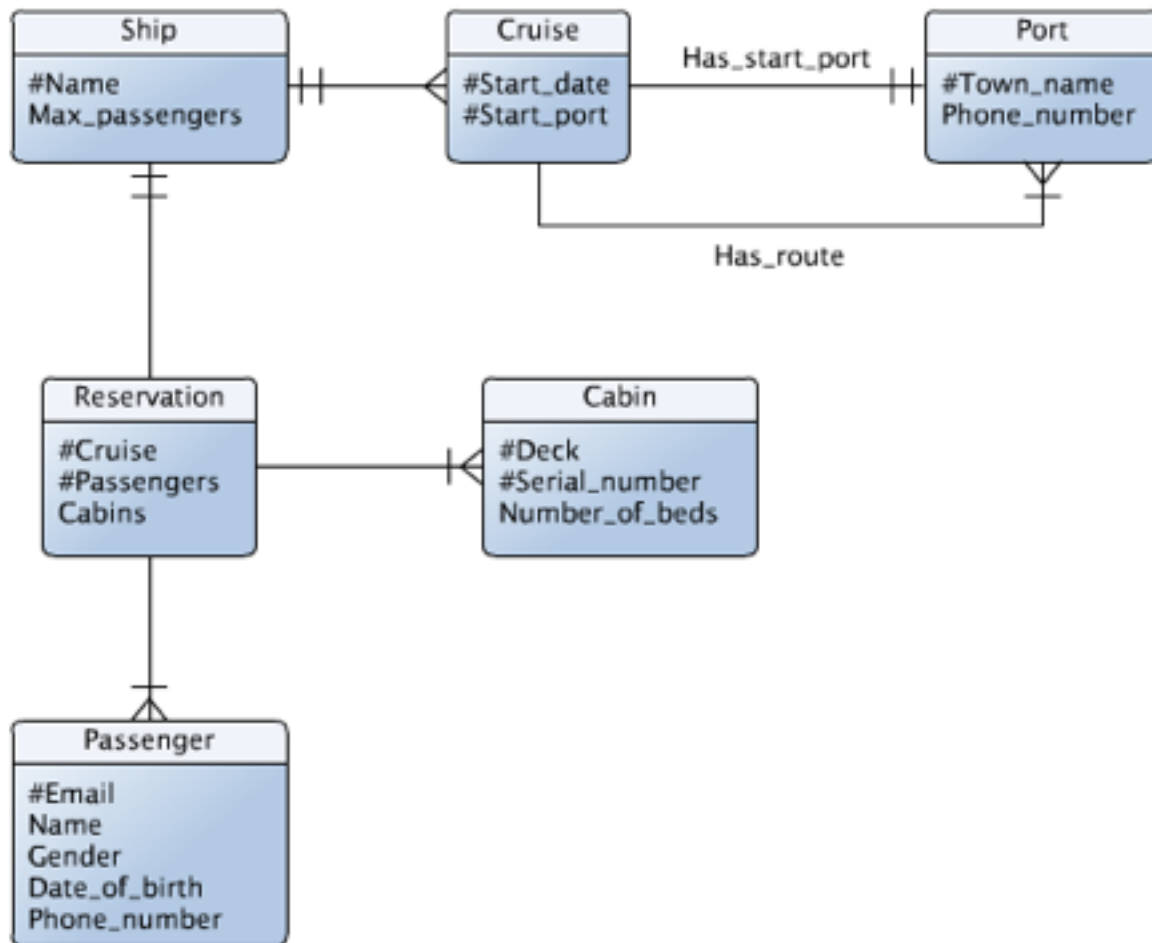
5i)
Truck is now a weak entity and has no primary key. One could for example register multiple trucks with the same registration number and different model/weight etc, which makes no sense.

ii)
Model -> Maximum_weight
Registration_number -> Registration_year

iii) {Registration_number}, {Registration_Number, Model}. The first one makes the most sense since having Model in the key introduces redundancy. Also having assignment_number in each truck makes no sense, since the truck exists independently of the assignment, but the assignment needs the truck. So the Assignment should have a foreign key to Truck instead.

iv)

Starting with the Truck table

Truck (Registration_number, Registration_year, Model, Maximum_weight, Assignment_number*)
->
Truck(#Registration_number, Registration_year, Model*)
Model(#Model, Maximum_weight)
Assignment(#Assignment_number, Telephone_number*, Container_number*, start_date, end_date, Truck*)

Truck* in Assignment is multi-value, so:
Assignment(#Assignment_number, Telephone_number*, Container_number*, start_date, end_date)
Truck_Assignment(Assignment_number*, Truck*)

An assignment can also have multiple containers involved so Container_number* is currently non-atomic:

Assignment(#Assignment_number, Telephone_number*, start_date, end_date)
Assignment_Containers(Assignment_number*, Container_number*)

Need to keep track of which containers are on which truck…
Truck_Container_Assignment(Assignment_number*, Truck*, Container_number*)

Final BCNF result:
Container_type (#Type_id)
Container (#Container_number, Type_id*)
Customer (#Telephone_number)
Assignment(#Assignment_number, Telephone_number*)
Truck_Container_Assignment(Assignment_number*, Truck*, Container_number*)
Truck(#Registration_number, Model*)
Model(#Model)

Note: Was not sure of the convention used for ER / logical table, while I haven't been to lectures it is unclear what it is compared to other conventions. For example in ER-diagrams here we see primary keys, in logical we see table names, and so on. Also the convention of having the <> figures to represent relationships seems to not be used in this course, instead just naming the relationships directly. Some sort of cheat-sheet or quick overview of exactly what convention is used in this course to represent ER diagrams, logical tables, cardinality and so on would be really helpful.(http://stackoverflow.com/questions/4279089/what-is-the-difference-between-logical-data-model-and-conceptual-data-model).

Feature	Conceptual	Logical	Physical
Entity Names	✓	✓	
Entity Relationships	✓	✓	
Attributes		✓	
Primary Keys		✓	✓
Foreign Keys		✓	✓
Table Names			✓
Column Names			✓
Column Data Types			✓