Assignment 1: Relational Database Normalization

작성자 2022-24670 송종현

0. Introduction

As the one who is interested in aerospace engineering and industry, data with regard to the Korean aviation industry were gathered. Downloaded CSV files are as below:

- Airworthiness* Directives (감항성개선지시현황) (Source)
- Korean Registered Aircrafts (항공기 등록현황) (Source)
- Aircraft Model Information (세계항공기 정보) (Source)
- Freight Performance on Incheon Int'l Airport (인천공항 항공사별 노선별 운송실적) (Source)

Airworthiness: the fact that an aircraft is in working condition and safe to fly

1. Preprocessing & DB Normalization

1-1 Airworthiness Directives (감항성개선지시현황) (Source)- 3 tables decomposed

AD 발행ë²^í~, (Number)	발행êµ-ê°€ (Issued by)	AD ē°œí-‰ì¼ìž (Issue date)	ì œëª© (Subject)	ice iš"i¼ (Effective date)	ë≢"ë, (Model)	ì™,êμ-AD ë²^í″,	i œiží),¬	i*,ê³ ìžë£Œ (Attachment)
			Repetitive inspections for cracking of					
			fastener holes at a certain station of the		B737-400			2020-26-09.pdf
2020-222	US	12/24/2020	center wing box	1/27/2021	B737-400	2020-26-09	BOEING COMPANY	AD ë~,ì,,œ
					AS-365N2			EASA_AD_2020-0287.pdf
			ATA 65 Tail Rotor Drive - Tail Rotor Drive		AS-365N2			AD ë-, i,, ce
			Flange / Shaft Flexible Coupling		AS-365N3			
020-174R1	EU	12/22/2020	Inspection / Modification	1/4/2021	EC-155B1	2020-0287	AIRBUS INDUSTRIE	
			Repetitive inspections for cracking of		B737-400			2020-26-04.pdf
013-189R1	US	12/22/2020	certain skin panels of the fuselage	1/25/2021	B737-400	2020-26-04	The Boeing Company	AD ë¬,ì,∞e
			ATA36 Pneumatic - Pylon / Wing					
			Interface Bleed Duct and Fuel Pipe					EASA_AD_2020-0286.pdf
020-140R1	EU	12/21/2020	Inspection	1/1/2021	A380-800	2020-0286	AIRBUS INDUSTRIE	AD ë-, ì,, œ
			ATA 64 Tail Rotor - Blade Re-					EASA_AD_2018-0168R1.pdf
2018-120R1	EU	12/21/2020	identification / Life Limit	12/25/2020	EC135P2+	2018-0168R1	AIRBUS HELICOPTERS	AD ë-, ì,, œ
			ATA 31 Indicating / Recording Systems -					
			Flight Data Recording System					EASA_AD_2020-0285.pdf
012-009R1	EU	12/21/2020	Modification (Software Update)	1/1/2021	A380-800	2020-0285	AIRBUS INDUSTRIE	AD ë¬_서
								EASA_AD_2020-0282.pdf
2020-221	EU	12/18/2020	ATA 64 Tail Rotor - Blades Inspection	12/31/2020	EC135P2+	2020-0282	AIRBUS HELICOPTERS	AD ë¬,ì₀œ
					A330-200			EASA_AD_2020-0283.pdf
2019-203R2	EU	12/18/2020	ATA 28 Fuel - Fuel Pump Inspection	12/31/2020	A330-300	2020-0283	AIRBUS INDUSTRIE	AD ë-,ì,œ
					GENX-2B67B			2020-25-10.pdf
					GENX-2B67/P			AD ë-,ì,œ
			Ultrasonic inspection (USI) of the HPT		GENX-1B			
2020-220	US	12/17/2020	rotor stage 2 disk	1/21/2021	GENX-1B	2020-25-10	GE	
	1		ATARES STATES AND A SECOND SEC			1		EACA AD 2020 0204 - 1/

There are several problems

- 1. It DOES NOT satisfy 1NF. The model column has multiple values.
- 2. The attachment column does not give any necessary information.

By preprocessing (deleting attachment column and giving unique value per each row), the given table has satisfied 1NF.

	asc Number ∜↓	ABC Issued by 🏋‡	ABC Issue date 🏋‡	ABC subject	ABC Effective date T:	ABC model 📆	ABC Foreign AD 17:	ABC manufacture T↓
1	2022-045	US	3/20/2022	ATA 34 Navigation - Airplane Flight Manual Revision	3/16/2022	B747-400	6/16/2022	The Boeing Company
2	2022-045	US	3/20/2022	ATA 34 Navigation - Airplane Flight Manual Revision	3/16/2022	B747-400(BDSF)	6/16/2022	The Boeing Company
3	2022-045	US	3/20/2022	ATA 34 Navigation - Airplane Flight Manual Revision	3/16/2022	B747-400F	6/16/2022	The Boeing Company
4	2022-045	US	3/20/2022	ATA 34 Navigation - Airplane Flight Manual Revision	3/16/2022	B747-400SF	6/16/2022	The Boeing Company
5	2022-045	US	3/20/2022	ATA 34 Navigation - Airplane Flight Manual Revision	3/16/2022	B747-400	6/16/2022	The Boeing Company
6	2022-045	US	3/20/2022	ATA 34 Navigation - Airplane Flight Manual Revision	3/16/2022	B747-400(BDSF)	6/16/2022	The Boeing Company
7	2022-045	US	3/20/2022	ATA 34 Navigation - Airplane Flight Manual Revision	3/16/2022	B747-400F	6/16/2022	The Boeing Company
8	2022-045	US	3/20/2022	ATA 34 Navigation - Airplane Flight Manual Revision	3/16/2022	B747-400SF	6/16/2022	The Boeing Company
9	2022-044	US	3/14/2022	ATA 54 Nacelles / Pylons - Inlet Cowl Modification	4/15/2022	B777-200	6/11/2022	The Boeing Company
10	2022-044	US	3/14/2022	ATA 54 Nacelles / Pylons - Inlet Cowl Modification	4/15/2022	B777-300	6/11/2022	The Boeing Company
11	2022-044	US	3/14/2022	ATA 54 Nacelles / Pylons - Inlet Cowl Modification	4/15/2022	B777-200	6/11/2022	The Boeing Company
12	2022-044	US	3/14/2022	ATA 54 Nacelles / Pylons - Inlet Cowl Modification	4/15/2022	B777-300	6/11/2022	The Boeing Company

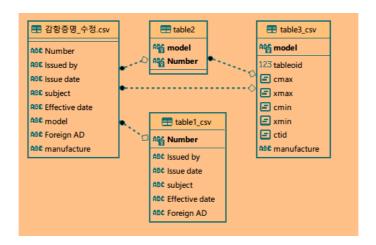
Let us check the relation between columns to normalize the given database to 2NF and 3NF.

Given the database in 1NF, we can figure out a tuple of (model, number) that can uniquely define each row. Thus, to be in 2NF, there should be no non-prime attribute that is functionally dependent on any subset of primary keys.

As seen above, If we know the number, we can know "Issued by", "Issue date", "subject", "Effective date", "Foreign AD". All of those are defined iff the number is given. Also, we can know manufacture when the model is given (ex. Boeing 777 is

manufactured by Boeing, literally)

Also, as there are no transitive dependencies between columns, the third normal form is completed as below.



(NOTE: unwanted columns revealed on table 3. the only intended is manufacture) model and number is a foreign key to table1 and table3, each.



(Model, Number): Primary Keys and foreign keys to table 1 and 3

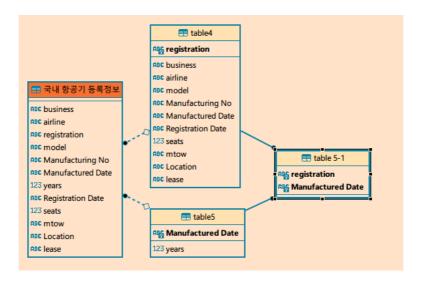
Model ⇒ manufacture

Number ⇒ (Issued by, Issue date, subject, Effective date, Foreign AD)

1-2 Korean Registered Aircrafts (항공기 등록현황)- 3 tables created



Unlike the first table, the "one value per each row" principle has been observed, the only dependence exists between "Manufactured date" and "years" and other columns functionally depend on registration, a code unique to a single aircraft.



Raw data has been split into two tables, whose primary keys are registration* and manufactured date.

(NOTE: dependencies between model, seats, and mtow-maximum takeoff weight- has been expected. However, depending on various options when purchasing, there were no such relationships between them.)



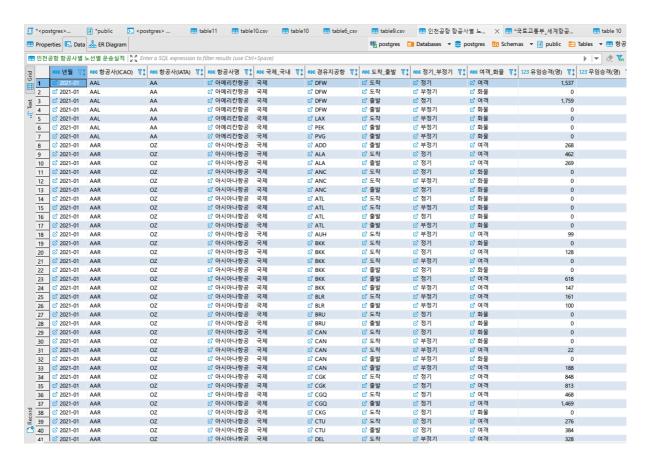
Summary

(Registration, Manufactured Date): Primary Keys and foreign key to table 4 and 5-1

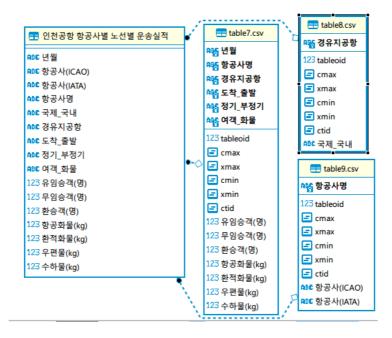
Registration (P key) ⇒ (business, airline, model, Manufacturing No, Manufactured Date, Registration Date, seats, mtow, Location, lease)

Manufactured Date (P key) ⇒ years

1-3 Freight Performance on Incheon Int'l Airport (인천공항 항공사별 노선별 운송실적)- 3 tables decomposed



The Database above shows overall transportation in Incheon int'l airport in 2021. To uniquely determine transported passengers and freight, 6 primary keys are required-년월, 항공사명, 경유지공항, 도착_출발, 정기_부정기, 여객_화물. Any other information-국제_국내 and ICAO & IATA code- is determined solely by 경유지공항 and 항공사명. To remove partial dependence (That is, to satisfy 2NF), Raw DB has been split into 3 tables whose primary keys are as below



(NOTE: unwanted columns revealed on every table)



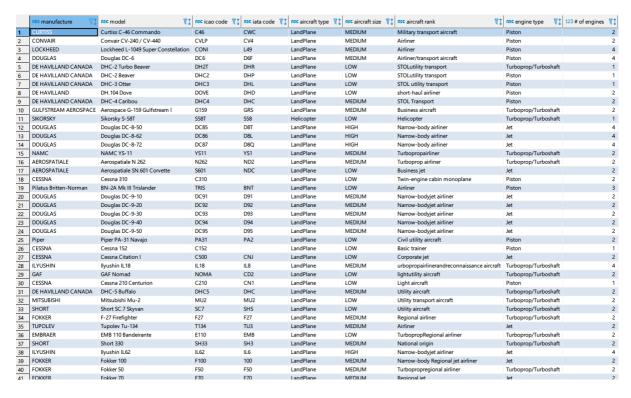
(년월, 항공사명, 경유지공항, 도착_출발, 정기_부정기, 여객_화물): Primary Keys and 경유지공항, 항공사명 are foreign key to table8 and table9

(년월, 항공사명, 경유지공항, 도착_출발, 정기_부정기, 여객_화물) **(P key)** ⇒ (유임승객, 무임승객, 환승객, 항공화물, 환적화물, 우편물, 수하물)

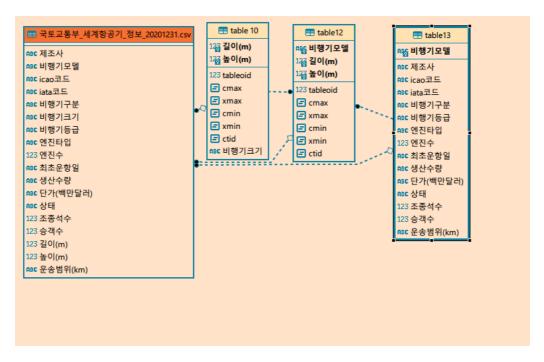
경유지공항 (P key) ⇒ 국제_국내

항공사명 (P key) ⇒ 항공사(ICAO), 항공사(IATA)

1-4 Aircraft Model Information (세계항공기 정보)- 3 Tables Decomposed



Like 1-2, it satisfies 1NF. Figuring out relations between columns, aircraft size may have an influence on length and height or vice versa. However, there was no explicit relationship. Surprisingly, the model can determine all the properties of an aircraft. (if not, model name lose its meaning) aircraft size(비행기 크기) can be determined by its height and length. Two databases whose primary keys are (길이, 높이) and 비행기모델 are normalized as below.



(NOTE: unwanted columns revealed on table 10 and 12)



Summary

(비행기모델, 길이, 높이): Primary Keys and foreign key to table 13 and 10 each.

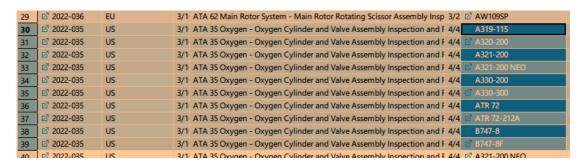
(길이, 높이) P-keys ⇒ 비행기 크기

비행기모델 P-key ⇒ 제조사,icao코드,iata코드,비행기구분,비행기등급,엔진타입,엔진수,최초운항일,생산수량,"단가(백만달러)",상태,조 종석수,승객수,"운송범위(km)"

2. Ten Interesting Questions

To make questions more *interesting*, let us imagine the situation where a public servant in aviation policy makes use of the databases above.

1. Airworthiness Directive 2022-035 "ATA 35 Oxygen - Oxygen Cylinder and Valve Assembly Inspection and Replacement" requires the airlines who own A319-115, A320-200, A321-200, A321-200 NEO. A330-200, A330-300, ATR 72, ATR 72-212A, B747-8F to report service bulletin within 15 days. Which companies in Korea are in charge of it?



```
select distinct model, airline from table4 t
where (model like '%A319-115%' or
model like '%A320-200%' or
model like '%A321-200%' or
model like '%A321-200 NEO%' or
model like '%A330-200%' or
model like '%A330-200%' or
model like '%A3430-200%' or
model like '%A370-300%' or
model like '%A377 72%' or
model like '%ATR 72-212A%' or
```

```
model like '%B747-8%' or
model like '%B747-8F%')
```

	asc model ∜‡	asc airline 🎵↑
1	A319-115	SK텔레콤
2	A330-200	대한항공
3	B747-8F	대한항공
4	B747-8	대한항공
5	A330-300	대한항공
6	A321-200	아시아나항공
7	A321-200 NEO	아시아나항공
8	A330-300	아시아나항공
9	A320-200	아시아나항공
10	A320-200	에어로케이항공
11	A320-200	에어부산
12	A321-200 NEO	에어부산
13	A321-200	에어부산
14	A321-200	에어서울
15	A330-300	티웨이항공
16	ATR 72-212A	하이에어
17	ATR 72	하이에어

SK Telecom, Korean Air, Asiana Air, Aero-K Air, Air Busan, Air Seoul, T-way air, hi-air are in charge of it

2. Delegates from the Gambia will visit S.Korea for economic purposes. The Korean government will lease the aircraft from private airlines. Distance between Gambia and S.Korea is 13,287km and 400 people will visit this time. which aircraft from which airline will be appropriate?

```
select 제조사, 비행기모델, "운송범위(km)", 승객수 from table6
where "운송범위(km)">13287 and 승객수>400;
// The first query gave A-330-200, A-330-800neo, A-330-900neo, A-350-900, A-350-1000,
//A-380-800, Boeing 747-400, Boeing 747-8I, Boeing 777-9, Boeing 777-300ER, Boeing 787-9
// as an answer
select distinct model, airline from table4 t
where (model like '%A330-200%' or
model like '%A330-800neo%' or
model like '%A330-900neo%' or
model like '%A350-900%' or
model like '%A350-1000%' or
model like '%A330-300%' or
model like '%A380-800%' or
model like '%B747-400%' or
model like '%B747-8I%' or
model like '%B777-9%' or
model like '%B777-300ER%' or
model like '%B787-9%')
```

Query above gave the answer that

```
B747-400(BDSF) 아시아나항공
B747-400F 아시아나항공
A330-200
         대한항공
         아시아나항공
A330-300
         아시아나항공
A350-900
         아시아나항공
A380-800
A380-800
         대한항공
B747-400F
         대한항공
         티웨이항공
A330-300
B747-400
         아시아나항공
B787-9 대한항공
A330-300
         대한항공
B777-300ER 대한항공
B747-400SF 아시아나항공
B787-9 에어프레미아 are appropriate.
```

3. A Big typhoon is forecasted in Jolla Province (전라도) Emergency aid should be carried out to prevent damage to aircraft. which airlines should be in contact in this situation?

```
select distinct "Location", airline from table4 t
where ("Location" like '전%') or ("Location" like '무안공항') or ("Location" like '광주공항')
```

전라남도, 트랜스헬리, 경운대학교 et cetera should be under appropriate measurement for typhoon

	ABC Location T:	ABC airline 🏋 🗘		
1	전남 영암군 미암면	(주)신한에어		
2	무안공항	써니항공		
3	전북임실관촌	트랜스헬리		
4	무안공항	숭선학원		
5	전남화순	중앙119구조본부		
6	전남 영암	경운대학교		
7	무안공항	(주)지오스토리		
8	무안공항	초당학원		
9	무안공항	한국교통대학교		
10	전남 영암군 덕진면 소방항공대길 95	산림청		
11	전남 영암	(주)신한에어		
12	전남 영암군	오**(1221)		
13	전남 영암군	경운대학교		
14	광주공항	아시아나항공		
15	전북 군산시 임피면 영창리 299-35	수에어(남철)		
16	전남 영암군	플라잉타이거즈		
17	무안공항	중원대학교		
18	전남 영암군 덕진면 소방항공대길99	전라남도		
19	전북 익산	산림청		
20	무안공항	경운대학교		
21	전북 남원요천비행장	한양항공		
22	무안공항	학교법인 청석학원		
23	무안공항	제**(1357)		
24	전남 영암	산림청		
25	광주공항	광주광역시		
26	무안공항	하이에어		
27	광주공항	티웨이항공		
28	전남 영암	전라남도		

4. For aviation safety, the law was enacted that aircraft over 40 years is prohibited to fly over Korean territory. which aircraft will be the target of this law?

```
select table5."Manufactured Date", table4."Manufactured Date", table5.years, table4.registration from table5, table4 where table5."Manufactured Date"=table4."Manufactured Date" and table5.years>40
```

Overall 50 planes will be prohibited to fly

	🚌 Manufactured Date 🏻 🕽 🛊	ABC Manufactured Date 🏋 🕻	123 years 🏋 🔭	নদুর registration 🏋
35	80.11.14	80.11.14	41	HL1103
36	62.02.01	62.02.01	60	HL9621
37	77.02.01	77.02.01	45	HL9629
38	73.05.15	73.05.15	48	HL9662
39	63.07.01	63.07.01	58	HL9663
40	75.02.17	75.02.17	47	HL9678
41	76.12.15	76.12.15	45	HL9238
42	80.06.10	80.06.10	41	HL1192
43	74.07.30	74.07.30	47	HL1137
44	66.10.06	66.10.06	55	HL0305
45	78.03.24	78.03.24	44	HL8373
46	80.03.20	80.03.20	42	HL9154
47	80.09.20	80.09.20	41	HL9159
48	79.03.12	79.03.12	43	HL9146
49	80.05.20	80.05.20	41	HL9158
50	80.10.12	80.10.12	41	HL9189

5. Korean government bestows *great trade awards* to the city whose freight performance is the best among the destinations of the flight from Incheon international airport. which city will be awarded in 2021?

```
select 경유지공항, SUM("항공화물(kg)")+SUM("환적화물(kg)")+SUM("우편물(kg)")+SUM("수하물(kg)")
from "table7.csv"
group by 경유지공항
order by SUM("항공화물(kg)")+SUM("환적화물(kg)")+SUM("우편물(kg)")+SUM("수하물(kg)") DESC
limit 6
```

Shanghai (PVG), Hong Kong (HKG), Los Angeles (LAX), Tokyo (NRT, HAN), Chicago (ORD) are honored.



The last half questions are for the aircraft manager of Korean Air.

6. Korean Air is considering purchasing new aircraft. The first flying date(최초운항일) should not be shall not exceed 10 years. The minimum # of seats should be over 200 for minimum profit. The cost per plane should be less than \$500M

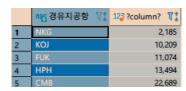
```
select 비행기모델, 최초운항일, 승객수, "단가(백만달러)" from table6 t
where 최초운항일 >'2012-03-28' and 승객수>200 and "단가(백만달러)"<500
```

A-321neo, A-330-800neo, A-330-900neo, A-350-900, A-350-1000, Boeing 777-8 could be possible options



7. Korean Air wants to cut 5 worst cities depending on freight performance to cut the cost. which are they?

Contract with Nanjing(NKG), Kagoshima (KOJ), Fukuoka (FUK), Cat Bi (HPH), bandaranaike (CMB) will be ended soon.



8. The aircraft model which Airworthiness Directive has required to fix most often will be the main topic for educating engineers. which was it?

```
select model, count(model) cnt
from table2 t
group by model
order by cnt desc
limit 1
```

A 320-200 will be the main topic



9. Ms. Cho, the owner of the company, wants to go to Machu Picchu, which is 16,500km distant from Seoul. As she is a very noble person, she wants to depart there at once, without transfer. Is it possible with any aircraft Korean air owns?

```
select table6.비행기모델, table6."운송범위(km)"
from table6
where "운송범위(km)">16500;
// The first query gave "A-340-500" only as an answer
select airline, model
from table4
where (airline like '%대한항공%')
// There is no "A-340-500" on the list
```

There is no aircraft available to the company.

10. Korean Air will add extra personnel to the cities to which the company actively transports passengers to enhance the passenger experience. The best 5 cities based on data in 2021 will be the ones. which are they?

```
select 경유지공항, SUM(case when 항공사명 = '대한항공' then "유임승객(명)" END)+SUM(case when 항공사명 = '대한항공' then "환승객(명)" END) sum
from "table7.csv"
group by 경유지공항
having SUM(case when 항공사명 = '대한항공' then "유임승객(명)" END)+SUM(case when 항공사명 = '대한항공' then "환승객(명)" END) is not null
order by SUM(case when 항공사명 = '대한항공' then "유임승객(명)" END)+SUM(case when 항공사명 = '대한항공' then "환승객(명)" END) DESC
limit 5
```

Extra personnel will be sent to Los Angeles, Atlanta, New York, Manila, Tokyo

