

Week 12 Final Project Diary

Wong Zi Xin

2023-11-20

Week 9 Diary

1. What is the topic that you have finalised?

The final topic that I have chosen is to create a data story and analyse air transportation data and information in the region of Europe. In the context of my analysis, the definition of Europe refers to the countries in the European Union to date: EU-27 countries. In particular, I will look at EU-27 countries with the most number of passengers and analyse which regions do most people travel to, and observe trends over the years. I chose this topic as I am very interested in travelling, and am curious to derive insights about how the air transport industry was impacted and has picked up again in a post COVID world.

2. What are the data sources that you have curated so far?

I mainly curated my data sources from Kaggle, using datasets containing information about the number of air transport passengers carried by country and datasets containing information on the list of airports and airlines globally.

Global datasets:

<https://www.kaggle.com/datasets/tjkyner/global-air-transport-data>

<https://www.kaggle.com/datasets/thedevastator/global-air-transportation-network-mapping-the-wo>

<https://www.kaggle.com/datasets/johnmwega/trends-and-insights-of-global-tourism>

Datasets specifically looking at Europe:

<https://www.kaggle.com/datasets/gpreda/passengers-air-transport-in-europe>

<https://data.europa.eu/data/datasets/38mt9yvqp2fhg7wwgqf13q?locale=en>

Week 10 Diary

1. What is the question you are going to answer?

How have air transportation trends in the EU-27 changed over time?

2. Why is this an important question?

According to the International Air Transport Association (IATA), air travel is one of the most important modes of transportation as the aviation industry contributes significantly to global GDP by facilitating global trade, business, tourism and more. With the outbreak of the COVID-19 pandemic IATA revealed the aviation industry suffered a loss of \$118 billion in 2020, but with the gradual revival of air travel post-COVID, insights into air travel can better inform strategies to foster economic recovery. Europe was selected as the focus region as according to the United Nations World Tourism Organisation (UNWTO), Europe is the world's top tourist destination.

Sources: <https://www.iata.org/en/ata-repository/publications/economic-reports/aviation-economic-benefits/>

<https://www.iata.org/en/ata-repository/publications/economic-reports/understanding-the-pandemics-impact-on-the-aviation-value-chain/>

<https://www.unwto.org/impact-assessment-of-the-covid-19-outbreak-on-international-tourism>

3. Which rows and columns of the dataset will be used to answer this question?

Columns that are useful to answer this question will be geo (for the country's name), TIME_PERIOD (to represent the corresponding year) and OBS_VALUE (to represent the total number of passengers). All rows are useful as they represent unique data entries of each country by year.

europe air passenger data 2022												
DATAFLOW	LAST UPDATE	freq	freq	freq	freq	freq	freq	freq	freq	freq	freq	freq
ESTAT:TTFR00012(1.0)	16/10/23 11:00:00	A	PAS	PAS_CRD	TOTAL	TOT	AT	2011	25137612			
ESTAT:TTFR00012(1.0)	16/10/23 11:00:00	A	PAS	PAS_CRD	TOTAL	TOT	AT	2012	25965977			
ESTAT:TTFR00012(1.0)	16/10/23 11:00:00	A	PAS	PAS_CRD	TOTAL	TOT	AT	2013	25749724			
ESTAT:TTFR00012(1.0)	16/10/23 11:00:00	A	PAS	PAS_CRD	TOTAL	TOT	AT	2014	243378676			
ESTAT:TTFR00012(1.0)	16/10/23 11:00:00	A	PAS	PAS_CRD	TOTAL	TOT	AT	2015	26754007			
ESTAT:TTFR00012(1.0)	16/10/23 11:00:00	A	PAS	PAS_CRD	TOTAL	TOT	AT	2016	27181511			
ESTAT:TTFR00012(1.0)	16/10/23 11:00:00	A	PAS	PAS_CRD	TOTAL	TOT	AT	2017	28327279			
ESTAT:TTFR00012(1.0)	16/10/23 11:00:00	A	PAS	PAS_CRD	TOTAL	TOT	AT	2018	31138417			
ESTAT:TTFR00012(1.0)	16/10/23 11:00:00	A	PAS	PAS_CRD	TOTAL	TOT	AT	2019	35644188			
ESTAT:TTFR00012(1.0)	16/10/23 11:00:00	A	PAS	PAS_CRD	TOTAL	TOT	AT	2020	9168431			
ESTAT:TTFR00012(1.0)	16/10/23 11:00:00	A	PAS	PAS_CRD	TOTAL	TOT	AT	2021	11105564			
ESTAT:TTFR00012(1.0)	16/10/23 11:00:00	A	PAS	PAS_CRD	TOTAL	TOT	AT	2022	25381180			
ESTAT:TTFR00012(1.0)	16/10/23 11:00:00	A	PAS	PAS_CRD	TOTAL	TOT	BA	2021	987659			
ESTAT:TTFR00012(1.0)	16/10/23 11:00:00	A	PAS	PAS_CRD	TOTAL	TOT	BE	2011	25102695			
ESTAT:TTFR00012(1.0)	16/10/23 11:00:00	A	PAS	PAS_CRD	TOTAL	TOT	BE	2012	25919515			
ESTAT:TTFR00012(1.0)	16/10/23 11:00:00	A	PAS	PAS_CRD	TOTAL	TOT	BE	2013	24338927			
ESTAT:TTFR00012(1.0)	16/10/23 11:00:00	A	PAS	PAS_CRD	TOTAL	TOT	BE	2014	28776258			
ESTAT:TTFR00012(1.0)	16/10/23 11:00:00	A	PAS	PAS_CRD	TOTAL	TOT	BE	2015	30958841			
ESTAT:TTFR00012(1.0)	16/10/23 11:00:00	A	PAS	PAS_CRD	TOTAL	TOT	BE	2016	30118832			
ESTAT:TTFR00012(1.0)	16/10/23 11:00:00	A	PAS	PAS_CRD	TOTAL	TOT	BE	2017	31630020			
ESTAT:TTFR00012(1.0)	16/10/23 11:00:00	A	PAS	PAS_CRD	TOTAL	TOT	BE	2018	34506309			
ESTAT:TTFR00012(1.0)	16/10/23 11:00:00	A	PAS	PAS_CRD	TOTAL	TOT	BE	2019	35385188			
ESTAT:TTFR00012(1.0)	16/10/23 11:00:00	A	PAS	PAS_CRD	TOTAL	TOT	BE	2020	9455828			
ESTAT:TTFR00012(1.0)	16/10/23 11:00:00	A	PAS	PAS_CRD	TOTAL	TOT	BE	2021	11630020			
ESTAT:TTFR00012(1.0)	16/10/23 11:00:00	A	PAS	PAS_CRD	TOTAL	TOT	BE	2022	27873892			
ESTAT:TTFR00012(1.0)	16/10/23 11:00:00	A	PAS	PAS_CRD	TOTAL	TOT	BG	2011	6652007			

Screenshot of a portion of dataset used

4. Include the challenges and errors that you faced and how you overcame them.

The main dataset that I am using, as seen in the screenshot provided when answering the above question, is not displayed in a very organised and tidy format. The data for each country for each year are all displayed as separate rows. This would make it difficult to create visualisation plots on R, therefore, the first thing I did was to tidy the dataset.

Firstly, I copied over only variables needed (geo, TIME_PERIOD and OBS_VALUE) into a new Excel sheet. Then, I used the pivot table function in Excel to reorganise the data to make it tidy.

Sum of OBS_VALUE Row Labels	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022 Grand Total
AT	25137612	25965977	25749724	26378676	26754007	27181511	28327279	31138417	35644188	9168431	11105564	26381180
BE	25102695	25919515	26389927	28776258	30958841	30115832	33260493	34506309	35385188	9455828	13000020	27873892
BG	6652007	7529102	7529102	7529102	7529102	7529102	7529102	7529102	7529102	7529102	7529102	9253404
CH	4143848	4329006	4421759	4612745	4803675	5005482	5268443	5613849	5718439	1600811	1918008	4043418
CY	7190387	7238900	7011733	7238546	7508787	8861817	11009651	12577701	11281410	2759077	5095074	8807421
DE	1265032	1174232	1189182	1207983	1267004	1367382	1645554	1783821	1876708	3821372	4755160	9325240
DK	17531676	17531676	17531676	17531676	17531676	17531676	17531676	17531676	17531676	17531676	17531676	18333443
EL	2580351	2652739	2745963	2901153	3006556	3276312	3328124	3470139	34780127	8696674	1081781	26754007
ES	180799	205427	185695	201808	216009	221499	236508	250500	265261	85763	129541	873137
EU	33770739	33062336	34023934	39117783	4209442	45543371	50170728	54558826	56088527	17341192	32245559	57893929
EU27_2020	160715230	159771261	157719123	162364382	174625203	182672327	203804565	220611429	225516232	57707750	91845011	102671203
FI	74266281	74046281	74010281	74010281	74010281	74010281	74010281	74010281	74010281	74010281	74010281	9124880025
FR	16374258	16458815	16555391	17171931	17479246	17696521	17763502	17763502	17763502	5428622	4554497	13812577
FR	126013257	126013257	126013257	126013257	126013257	126013257	126013257	126013257	126013257	126013257	126013257	149116046
HR	4898236	5425232	5722447	6140797	6571698	7475463	8843003	9731284	10623399	1943547	6846803	9415321
HU	8403837	8403837	8403837	8403837	8403837	8403837	8403837	8403837	8403837	8403837	8403837	122944042
IE	2336289	23584089	24063840	26310826	28545020	32589709	34271771	36345005	37847510	8288297	9097359	32405890
IT	11622687	11622687	11622687	11622687	11622687	11622687	11622687	11622687	11622687	11622687	11622687	13538202
LT	3508814	3508814	3508814	3508814	3508814	3508814	3508814	3508814	3508814	3508814	3508814	43761984
LU	5098390	4754530	4722527	4802282	5145856	5384160	6077854	7037070	7785728	1995133	2336134	5368369
MT	3508814	3508814	3508814	3508814	3508814	3508814	3508814	3508814	3508814	3508814	3508814	55473084
NL	3040329	34380225	36969364	37602195	37602195	37602195	37602195	37602195	37602195	37602195	37602195	38600009
NO	2063572	2180075	2274484	2571422	28907439	32266861	37684668	43767548	46842771	13825460	18838312	39347542
PT	27579707	28186254	29684146	32560211	36005814	47673057	51018598	55007894	56007894	16548993	22347692	57081723
RO	9687458	9674226	10016933	10907487	12580711	15153719	17934774	19809542	21656855	6535447	10346013	19333951
SE	16374398	16458815	16555391	17171931	17479246	18099954	20054947	22173330	23287929	5428622	4554497	13812577
FR	126013257	126013257	126013257	126013257	126013257	126013257	126013257	126013257	126013257	126013257	126013257	149116046
HR	4898236	5425232	5722447	6140797	6571698	7475463	8843003	9731284	10623399	1943547	6846803	9415321
HU	8403837	8403837	8403837	8403837	8403837	8403837	8403837	8403837	8403837	8403837	8403837	122944042
IE	2336289	23584089	24063840	26310826	28545020	32589709	34271771	36345005	37847510	8288297	9097359	32405890
IT	11622687	11622687	11622687	11622687	11622687	11622687	11622687	11622687	11622687	11622687	11622687	13538202
LT	3508814	3508814	3508814	3508814	3508814	3508814	3508814	3508814	3508814	3508814	3508814	43761984
LU	5098390	4754530	4722527	4802282	5145856	5384160	6077854	7037070	7785728	1995133	2336134	5368369
NL	3040329	34380225	36969364	37602195	37602195	37602195	37602195	37602195	37602195	37602195	37602195	38600009
NO	2063572	2180075	2274484	2571422	28907439	32266861	37684668	43767548	46842771	13825460	18838312	39347542
PT	27579707	28186254	29684146	32560211	36005814	47673057	51018598	55007894	56007894	16548993	22347692	57081723
RO	9687458	9674226	10016933	10907487	12580711	15153719	17934774	19809542	21656855	6535447	10346013	19333951
SE	16374398	16458815	16555391	17171931	17479246	18099954	20054947	22173330	23287929	5428622	4554497	13812577
FR	126013257	126013257	126013257	126013257	126013257	126013257	126013257	126013257	126013257	126013257	126013257	149116046
HR	4898236	5425232	5722447	6140797	6571698	7475463	8843003	9731284	10623399	1943547	6846803	9415321
HU	8403837	8403837	8403837	8403837	8403837	8403837	8403837	8403837	8403837	8403837	8403837	122944042
IE	2336289	23584089	24063840	26310826	28545020	32589709	34271771	36345005	37847510	8288297	9097359	32405890
IT	11622687	11622687	11622687	11622687	11622687	11622687	11622687	11622687	11622687	11622687	11622687	13538202
LT	3508814	3508814	3508814	3508814	3508814	3508814	3508814	3508814	3508814	3508814	3508814	43761984
LU	5098390	4754530	4722527	4802282	5145856	5384160	6077854	7037070	7785728	1995133	2336134	5368369
NL	3040329	34380225	36969364	37602195	37602195	37602195	37602195	37602195	37602195	37602195	37602195	38600009
NO	2063572	2180075	2274484	2571422	28907439	32266861	37684668	43767548	46842771	13825460	18838312	39347542
PT	27579707	28186254	29684146	32560211	36005814	47673057	51018598	55007894	56007894	16548993	22347692	57081723
RO	9687458	9674226	10016933	10907487	12580711	15153719	17934774	19809542	21656855	6535447	10346013	19333951
SE	16374398	16458815	16555391	17171931	17479246	18099954	20054947	22173330	23287929	5428622	4554497	13812577
FR	126013257	126013257	126013257	126013257	126013257	126013257	126013257	126013257	126013257	126013257	126013257	149116046
HR	4898236	5425232	5722447	6140797	6571698	7475463	8843003	9731284	10623399	1943547	6846803	9415321
HU	8403837	8403837	8403837	8403837	8403837	8403837	8403837	8403837	8403837	8403837	8403837	122944042
IE	2336289	23584089	24063840	26310826	28545020	32589709	34271771	36345005	37847510	8288297	9097359	32405890
IT	11622687	11622687	11622687	11622687	11622687	11622687	11622687	11622687	11622687	11622687	11622687	13538202
LT	3508814	3508814	3508814	3508814	3508814	3508814	3508814	3508814	3508814	3508814	3508814	43761984
LU	5098390	4754530	4722527	4802282	5145856	5384160	6077854	7037070	7785728	1995133	2336134	5368369
NL	3040329	34380225	36969364	37602195	37602195	37602195	37602195	37602195	37602195	37602195	37602195	38600009
NO	2063572	2180075	2274484	2571422	28907439	32266861	37684668	43767548	46842771	13825460	18838312	39347542
PT	27579707	28186254	29684146	32560211	36005814	47673057	51018598	55007894	56007894	16548993	22347692	57081723
RO	9687458	9674226	10016933	10907487	12580711	15153719	17934774	19809542	21656855	6535447	10346013	19333951
SE	16374398	16458815	16555391	17171931	17479246	18099954	20054947	22173330	23287929	5428622	4554497	13812577
FR	126013257	126013257	126013257	126013257	126013257	126013257	126013257	126013257	126013257	126013257	126013257	149116046
HR	4898236	5425232	5722447	6140797	6571698	7475463	8843003	9731284	10623399	1943547	6846803	9415321
HU	8403837	8403837	8403837	8403837	8403837	8403837	8403837	8403837	8403837	8403837	8403837	122944042
IE	2336289	23584089	24063840	26310826	28545020	32589709	34271771	36345005	37847510	8288297	9097359	32405890
IT	11622687	11622687	11622687	11622687	11622687	11622687	11622687	11622687	11622687	11622687	11622687	13538202
LT	3508814	3508814	3508814	3508814	3508814	3508814	3508814	3508814	3508814	3508814	3508814	43761984
LU	5098390	4754530	4722527	4802282	5145856	5384160	6077854	7037070	7785728	1995133	2336134	5368369
NL	3040329	34380225	36969364	37602195	37602195	37602195	37602195	37602195	37602195	37602195	37602195	38600009
NO	2063572	2180075	2274484	2571422	28907439	32266861	37684668	43767548	46842771	13825460	18838312	39347542
PT	27579707	28186254	29684146	32560211	36005814	47673057	51018598	55007894	56007894	16548993	22347692	57081723
RO	9687458	9674226	10016933	10907487	12580711	15153719	17934774	19809542	21656855	6535447	10346013	19333951
SE	16374398	16458815	16555391	17171931	17479246	18099954	20054947	22173330	23287929	5428622	4554497	13812577
FR	126013257	126013257	126013257	126013257	126013257	126013257	126013257	126013257	126013257	126013257	126013257	149116046
HR	4898236	5425232	5722447	6140797	6571698	7475463	8843003	9				