Introduction

You have recently started a new airline company in Europe. The hub of your airline is in a major European city. Apart from your main hub, there are 19 other European destinations available to operate flights to.

The goal of this assignment is 1) to develop a network and fleet plan and 2) to solve a crew pairing problem for your airline. Both problems and their results are independent.

The appendices attached to this assignment contain all the data required to determine the revenue, cost, etc. for the development of the network and fleet plan and crew pairings.

In particular, the Excel-file *Assignment1_Problem1_Datasheets.xlsx* contains data on the location of your hub, the available airports, the demand for each Origin-Destination pair, and population and GDP data for the different cities and countries, respectively.

For the crew pairing problem, each group is given three unique datasheets containing a 1) flight schedule, 2) duty periods and 3) crew hotel costs.

First read the entire assignment carefully (including the appendices) to extract all information required to adapt the models presented in the lectures!

Problem 1

A. Demand forecast

In the first part of the assignment, your job is to start up operations of your airline. In your starting year (2020), your home country only has Air Services Agreement (ASAs) with other European countries. The ASAs within Europe allow up to the 9th Freedom of the Air. Unfortunately, the demand data that your airline possesses origins from 2015 and is thus outdated. Also, the most recent GDP and Population data is from 2018. It is up to you to forecast the demand for 2020. This can be done by using the following gravity model:

$$D_{ij} = k\,\frac{\left(pop_i pop_j\right)^{b1}\left(GDP_i GDP_j\right)^{b2}}{\left(f \cdot d_{ij}\right)^{b3}}$$

You will assume there is no global pandemic or other external factor affecting the demand.

Where:

D_{ij} is the demand between airports i and j

pop_i, pop_j, GDP_i and GDP_j are respectively the population and GDP for airports i and j. (population is per 1000 inhabitants, GDP is in USD)

f is the fuel cost, which is assumed to remain constant between 2015 and 2020 [USD/gallon]

d_{ij} is the distance between airports i and j.

To achieve this goal, you need to:

1. Calibrate the gravity model (find scaling factor k and coefficients $b1$, $b2$ and $b3$) using the 2015 demand values that are given in data file *AE4423_Datasheets.xlsx.*
    a. Tip 1: apply logarithms to linearize the gravity model.
    b. Tip 2: use ordinary least squares to obtain the best fit for the resulting linear formula. You can do this using Excel, Python or Matlab toolboxes, or any statistics software.
2. Forecast the population and GDP for 2020, based on the data for 2015 and 2018. You can assume a linear variation.
3. Generate the future demand using the calibrate gravity model and the estimated values for the population and GDP.

B. Network & fleet development

Equipped with accurate demand data, it is up to you to generate the weekly flight frequency plan for your airline. As the airline is new, you must acquire new aircraft. Thus, you are also asked to determine how many aircraft should be leased to maximize the profit.

To achieve this goal, you need to:

1. Adapt the **mathematical model** discussed in Lecture 3 to solve the problem. Write it down the formulation on a piece of paper.
2. Setup a **computer model** (e.g., in Python or MATLAB using CPLEX or GUROBI, or any other software that you wish to use) according to the model written on paper.
3. With the data available in the Excel file supplied with the assignment and the Appendix, determine the **network to be operated and the corresponding flight frequency** assuming one standard week of operations. Also determine **how many aircraft of each type** will be leased in the most optimal solution.

Additional Information

- **Although a real European network will be used in this assignment, the input data is synthetic, so treat the results accordingly!**
- This problem only deals with **intra-European** flights and passengers. Carefully check the corresponding **yield** and **load factors** in the appendices.
- You can lease as much aircraft as you desire, if this maximizes profit.
- Assume that the aircraft are only available for operations for **10 hours** per day, ensuring sufficient time for maintenance and no operations in periods of the day with very low demand.
- Turn-Around-Times (TAT), including landing and take-off times, **depend on the aircraft type** and the route operated. To allow for good connections in the hub, it should be assumed that the TAT for flights **to the hub** are **50% longer** than the normal TAT. Typical TATs per aircraft type can be found in Table 1.
- To start coding the problem, first create a script that reads/gathers all required input data and performs the required pre-processing to obtain distances, yield, cost, etc.

Problem 2

Crew Pairing Problem

You airline is now ready to start operations. Before that, you need to prepare the crew schedule. In this question you are asked to solve a simplified version of the **crew pairing problem.** We will only consider solving the crew pairing problem for a single day and all possible pairing (or duty periods, given that you are dealing with pairings of one day) are provided to you.

Consider the set of .csv files given for your group. The file “1_Timetable.csv” presents the daily flight schedule for your airline. Given this flight schedule, a set of legal daily duty periods are also provided to you in the file “2_Duty_Periods.csv”. Each duty period has an identifier number and consists of a list of ordered flights for a crew team.

The objective is to select the duty periods that will minimize the cost for the airline while ensuring that all the flights are covered by a single duty period. To efficiently solve the problem, you need to use the **column generation algorithm** discussed during the lectures and follow the steps below:

1. Compute the **costs associated with each duty period**. All monetary units are expressed in monetary units (MU). Assume that the total cost associated with a duty period consists of **3 components**. The **first** is a fixed salary that your airline has to pay per day to the crew regardless of how long they are on duty. The **second** is a flight duty pay, being salary component based on their duty duration. The **third** component of the cost are overstay costs (= hotel costs) that your airline has to pay if the crews do not return to their home base at the end of the day.
   The make-up of the crew salary costs are given in appendix E “Crew Costs”. The method to compute the duty duration is given in the next page under “additional information”. The overstay costs depend on the location of the crew; the per night hotel room costs for various cities, which are given in “3_Hotel_Costs.csv”.
2. Adapt the crew pairing formulation from Lecture 6 for this problem. Solve the **initial crew pairing problem** by only including duty periods that consist of a single flight.
3. Compute the **pricing problem**, obtaining the slackness for the decision variables not included in the model.

4. **Add the decision variables** that can help to improve your current solution to your crew pairing problem But don't add more than 50 new decision variables (i.e., duty periods) per iteration
5. **Iteratively repeat steps 3 and 4** until the optimal objective value cannot be improved further.

As part of the answer to this question, please provide the following:

- the total costs associated with duty periods number 100, 500, 750, 1200 and 2000.
- The optimal objective value obtained for the initial crew pairing problem (i.e., only with the duty periods with a single flight).
- A graph showing how the objective function value varied per iteration. Comment on the shape, characteristics of the resulting graph. Also give the runtime of the column generation procedure.
- A comment on the final result, indicating what is the daily total crew cost to run this flight schedule and the type of pairings (number of flights, duration, overnight costs, etc.) involved in the final solution.
- The list of the 20 duty periods with the lowest identifier number.
- A breakdown of the number of crew that will be required at each base.

Additional Information

- The **home base** of a crew member is the same as their starting base for each duty period.
- Assume the **duty period duration** to be the sum of the duty durations of all the flights in the duty period. The duty duration of a flight is equal to its flight time plus 2 brief periods. These brief periods represent the time the crew must spend finalizing procedures before/after a flight and their duration is provided in the input files.
- Each crew member will get his/her own hotel room in case of an overnight stay. The number of crew members per duty period is provided in Appendix E.
- To avoid unnecessarily long computation times, you can end the column generation procedure if the absolute changes in the objective function for the last 6 iterations **are all less than 0.001 MU**.

General information

- **Motivate your choices, comment on results, and be critical towards results!**

- Describe the mathematical model, your assumptions, results and KPIs for each of the assignments in detail in a comprehensive report of **no more than 15 pages A4** (excluding cover but including appendixes; font equivalent to Times New Roman – 12 pt., line spacing 1.15 and standard margins). **Note that the report shall not contain any computer code.**
- Use **figures** and **tables** to present your results and KPIs and support your conclusions.
- Submit your report and model script file(s) through **BrightSpace** (assignment folder in our course webpage) at the **latest on Friday 18 December, 18.00 hrs**. Do not forget to include the group number, names, and student IDs in the report (and script file(s)). Do **NOT** submit input (Excel) files. Files submitted by email will not be considered.
- **If you fail to meet the deadline**, 0.5 points will be deducted from your grade for **each day** after the deadline. No excuses will be accepted! Make sure that you work as a group and save the latest versions of your work in multiple places.
- All files uploaded in BrightSpace should be uploaded as **individual files** (i.e., do not compressed as '.zip', '.tar',…) to be subjected to **Turnitin check**. **If compressed files are uploaded**, 1.0 points will be deducted from your grade.
- **If you fail to obtain a grade of 5.5 or higher you will fail the assignment.** In that case, you will get a chance to improve your work and pass the assignment. Your final grade cannot become higher than 6.0 in that case.
- You should include a separate overview of the **workload distribution** of each group member. Indicate (in percentages) each member's contribution to the three categories **mathematical modelling (30%)**, **programming (50%)** and **reporting (20%)**. Based on this overview you will receive an individualized grade for the assignment. For an example of the format see Appendix F.
- **An assessment matrix** will be available to clarify the grading process.

Appendices

A. Revenue

Revenue is generated by transporting passengers. To determine the revenue for Problem 1, yield is expressed in € per Revenue-Passenger-Kilometer (RPK). The following revenue formula is to be used:

– The yield depends on the distance and can be formulated as follows:

$$Y_{EUR_{i,j}} = 5.9 \cdot d_{ij}^{-0.76} + 0.043$$

where:

- $Y_{EUR_{i,j}}$ is the yield in € between origin i and destination j.
- d_{ij} is the distance in km between origin i and destination j.

To develop a profitable network, fleet and frequency plan, the average load factor needs to be estimated. You can assume a load factor of 80% for this assignment.

B. Costs

Operating your aircraft logically induces costs. For Problem 1 of this assignment, two types of cost need to be considered:

- All aircraft are leased, and therefore a leasing cost needs to be accounted for. The weekly leasing cost is a fixed amount depending on the type of aircraft, and can be found in Appendix D, Table 1.
- Operating costs consist of three components:
  - Fixed operating costs (C_X^k) are costs **per flight leg** and represent costs such as landing rights, parking fees and fixed fuel cost. They depend only on the aircraft type k.
  - Time-based costs (C_T^k) are costs that are defined in € per flight hour and represent time-dependent operating costs such as cabin and flight crew. They depend on the distance of the flight leg and the aircraft type k. Time costs can be defined as follows:

    $$C_{T_{ij}}^k = c_T^k \frac{d_{ij}}{V^k}$$

    where:
    - $C_{T_{ij}}^k$ is the total time cost for a flight leg between airports i and j, operated by aircraft type k.
    - c_T^k is the time cost parameter for aircraft type k.
    - V^k is the airspeed of aircraft type k.
  - Fuel costs (C_F^k) are dependent on the distance flown, and can be expressed as follows:

    $$C_{F_{ij}}^k = \frac{c_F^k \cdot f}{1.5} d_{ij}$$

    where:
    - $C_{F_{ij}}^k$ is the fuel cost for a flight leg between airports i and j, operated by aircraft type k.
    - c_F^k is the fuel cost parameter for aircraft type k.
    - f is the fuel cost, equal to 1.42 USD/gallon in 2020.

The total operating cost for a flight leg between airports i and j, operated by aircraft type k can then be expressed as:

$$C_{ij}^k = C_X^k + C_{T_{ij}}^k + C_{F_{ij}}^k$$

The required parameters to determine the operating costs can be found in Table 1.

It should be noted that for flights departing or arriving at your hub airport all operating costs can be assumed to be 30% lower due to

economies of scale. (this includes fixed operating costs, time-based costs, and fuel costs).

C. Airport data

The data available in the provided Excel-sheet contains, among others, the position of each of the 20 available airports expressed in latitude (φ) and longitude (λ). The distance between two airports i and j can be determined using the following equations:

$$\Delta\sigma_{ij} = 2 \arcsin \sqrt{\sin^2\left(\frac{\varphi_i - \varphi_j}{2}\right) + \cos\varphi_i \cos\varphi_j \sin^2\left(\frac{\lambda_i - \lambda_j}{2}\right)}$$

$$d_{ij} = R_E \Delta\sigma_{ij}$$

where:

1. $\Delta\sigma_{ij}$ is the arc length between airports i and j.
2. $\varphi_i, \varphi_j, \lambda_i$ and λ_j are, respectively, the latitude and longitude for airports i and j.
3. R_E is the radius of the Earth in km; this can be assumed to be 6371 km.

Note that you can use online tools to determine the great circle distance between two locations on the Earth to confirm the correct implementation in your code.

Finally, you can assume there are no restrictions in the number of airport slots you can use.

D. Aircraft data

You have the choice to operate any of the four aircraft defined in Table 1.

Table 1: Aircraft data

Aircraft type	**Aircraft 1**: Regional turboprop	**Aircraft 2**: Regional jet	**Aircraft 3**: Single aisle twin engine jet	**Aircraft 4**: Twin aisle, twin engine jet
Aircraft characteristics				
Speed [km/h]	550	820	850	870
Seats	45	70	150	320
Average TAT [mins]	25	35	45	60
Maximum range [km]	1,500	3,300	6,300	12,000
Runway required [m]	1,400	1,600	1,800	2,600
Cost				
Weekly lease cost [€]	15,000	34,000	80,000	190,000
Fixed operating cost C_X [€]	300	600	1250	2000
Time cost parameter C_T [€/hr.]	750	775	1400	2800
Fuel cost parameter C_F	1.0	2.0	3.75	9.0

E. Crew Costs

The fixed salary and the flight duty pay of the various crew members are given in the table below.

Crew Type	Quantity	Fixed Salary [MU/day]	Flight Duty Pay [MU/hr]
Captain	1	98	120
First Officer	1	35	55
Steward(ess)	3	15	18

F. Individual workload

To distinguish between each student's workload in the group, you are required to provide an indication of each group member's workload in three separate disciplines. Provide the workload distribution in a separate file uploaded along with the assignment and follow the template below (or similar).

Student names	Mathematical modelling (30%)	Programming (50%)	Reporting (20%)
Student name #1	# %	# %	# %
Student name #2	# %	# %	# %
Student name #3	# %	# %	# %