

## Melton Scholarship Mini-Case (updated March 13, 2021)

### Overview of mini-case

The overall goal of this mini-case is for you to get a taste for what “doing guided research” looks and feels like. Like being in a class, you’ll be asked to perform weekly tasks with your mentor, “due” the following week when you meet up again and discuss the results. However, these tasks aren’t like homework problems you are assigned in classes. Typically, they are much more open-ended, require a substantial bit of thought and trial and error to solve, and they might not even have any “right” answers (e.g., when you’re exploring relationships in data find finding ones that are interesting). With that in mind, this mini-case would be along the lines of something you’d expect to do during a week!

### Background

The subject of this mini-case is airplane delays. The Bureau of Transportation keeps an online, “queryable” database of nearly all aspects of every US flight, including delays.

[https://www.transtats.bts.gov/Fields.asp?gnoyr\\_VQ=FGJ](https://www.transtats.bts.gov/Fields.asp?gnoyr_VQ=FGJ)

Click the “Download” on the list of options on the left-side of the page.

Some selectable fields that are interesting are:

- FlightDate (or other temporal field)
- Reporting\_Airline (see lookup table) – name of airline, like Delta, United, etc.
- OriginAirportID – name of airport that airline is flying out of, like TYS (Knoxville) and ORD (Chicago)
- DestAirportID – name of airport that airline is flying into
- CRSDepTime – scheduled departure time
- DepTime – actual departure time
- DepDelayMinutes – difference in minutes between scheduled and actual departure times
- DepDel15 – 1 = delayed by 15 or more minutes, 0 otherwise
- CRSArrTime – scheduled arrival time
- ArrTime – actual arrival time
- ArrDelayMinutes – difference in minutes between scheduled and actual arrival times
- ArrDel15 - 1 = delayed by 15 or more minutes, 0 otherwise
- Cancelled – 1 = yes flight was canceled, 0 otherwise; note the Code column too!
- WeatherDelay – delay due to weather, in minutes

Chicago O'Hare (ORD) is one of the nation's biggest airports. In winter months, major weather events have the possibility of wreaking havoc and severely delaying or cancelling flights. How big of an impact do major events have? How do delays that originate in O'Hare propagate across the country and influence delays elsewhere? Imagine that your faculty mentor has asked you to explore these questions (see what I mean by being open ended?).

As part of your application to the Melton Scholar program, you will include a short write-up that addresses these questions. We are interested in seeing how you think, i.e., how you determined which data you needed to collect, what you decided to measure, what numbers you decided to calculate and compare, what graphics you thought could be useful! Thus, your writeup isn't judged based on the absolute correctness of the final results (these are always discussed and modified by the faculty mentor). Rather, it will be judged based on your approach to the problem and what you decided to look at.

If this were actually your project, you'd take a stab at it and report to your mentor what you found, then you and your mentor would have a dialog over the implications of your work, what might be improved, and what should be looked at next. Research is very much an interactive and iterative process! The "final" results often take weeks or months of going back to the drawing board, trying out a different approach, and gradually understanding more and more about the problem until the most effective line of analysis seems clear.

### **Guidelines:**

- No minimum length, but try to summarize everything in no more than 3-4 pages.
- The writeup should give a narration of your approach and what you did (since we want to see how you think), not JUST the results (we want to understand where they came from).
- The case as written leaves a lot of room for exploration. You don't have to consider *everything* your "mentor" has asked about, but you should present something relevant to at least something your "mentor" mentioned!
- Some numbers would be nice. A well-designed graphic would be nicer.
- No restriction on what type of software to use. You can just use Excel if that's what you would like. R, Tableau, Python, etc., can do some neat things but it's not required to use them.
- Feel free to ask Prof. Petrie ([apetrie@utk.edu](mailto:apetrie@utk.edu)) questions and to "pretend" like he is your mentor.
- Submit a Word document (perhaps saved as a PDF) to Prof. Petrie after your application has been submitted, and by no later than the date specified on the Melton Scholar website.

## Some minimal guidance:

Most of the time, tackling a big, open ended problem like this involves studying just a small part of the picture at first, then slowly chipping away at a bigger problem. So coming up with a smaller, more easily approachable problem is a good place to start!

- The linked website allows all information about a particular month. It would be nice to have more data at once, but you might have to download a few months worth of data manually.
- How can you figure out what days had “major events”? Historical weather data is remarkably hard to find. It would be nice to find a site that gives daily snowfall totals in Chicago or ORD, but that might not exist.
- The site  
[https://www.weather.gov/lot/201617\\_winterevents](https://www.weather.gov/lot/201617_winterevents)

lists a few major winter events in Chicago for the 2016-2017 year, and maybe similar pages exist.

- The site:

<https://www.wunderground.com/weather/us/il/chicago-o-hare-international>

has “History” tab, and you can select Monthly on the page that comes up which summarizes each day of the month, events, and rain/snowfall totals. Worst case scenario you could scan through Nov-March of the last 5 years and see when the major events took place.

- Since you’d want to do something like compare the average departure delay out of ORD on days with major events to days with no major events, you’d want to also consider a few “normal” days too, so the weather data would come in handy again.
- Since Dec 4, 2016 had the first event of the season, had heavy wet snow, and was the largest December snowfall at ORD in 11 years, perhaps that’s a good date to look at. Maybe compare what happened that date with a date nearby that didn’t look to have any weather events.
- Think about where you can take it from here.