C.J.’s notes

* ASRP started 4/30/1975, modified in 1983, 1987, 1999 ([section 6](https://asrs.arc.nasa.gov/overview/immunity.html))
* Data are [anonymized](https://asrs.arc.nasa.gov/overview/confidentiality.html), FAA does not use data against reporters in enforcement (except for accidents and criminal offenses – cf. [FAR 91.25](https://asrs.arc.nasa.gov/overview/immunity.html))
* Description of [airspace classification](https://aspm.faa.gov/aspmhelp/index/Airspace_Classification.html)
* [Coding](https://asrs.arc.nasa.gov/docs/dbol/ASRS_CodingTaxonomy.pdf) Taxonomy, may be useful in interpreting database entries
* There will likely be some big effects from covid. (How) do we need to address that?
* Should we be looking at safety in accidents per flight or accidents per hour / mile?
  + I guess we can see what the data have to say here
  + Presumably takeoff and landing are the most dangerous segments of a flight, in which case accidents/flight makes more sense
* Jake’s proposed problems
  + NTSB – in the universe of aviation accidents, what are the predictors of severity?
  + ASRS – determine prominence of various types of incidents over time

5/27 -- Questions we have

* What is the research question?
* What data do we need to withhold for testing?

Tthings to share at meeting on 6/10

* Large dataset has number of uninjured passengers, so we can determine how many people were onboard
  + But there are a good handful (~2%) where this number is greater than the number of seats, in some cases by a wide margin
* Is there a way using T-100 queries to sort by flight type? (Re: seasonality of crashes / 100k flights)
* In ~5/6 fatal flights, everyone onboard died
* We impute values after train/test split, right?
* I did a train/test split