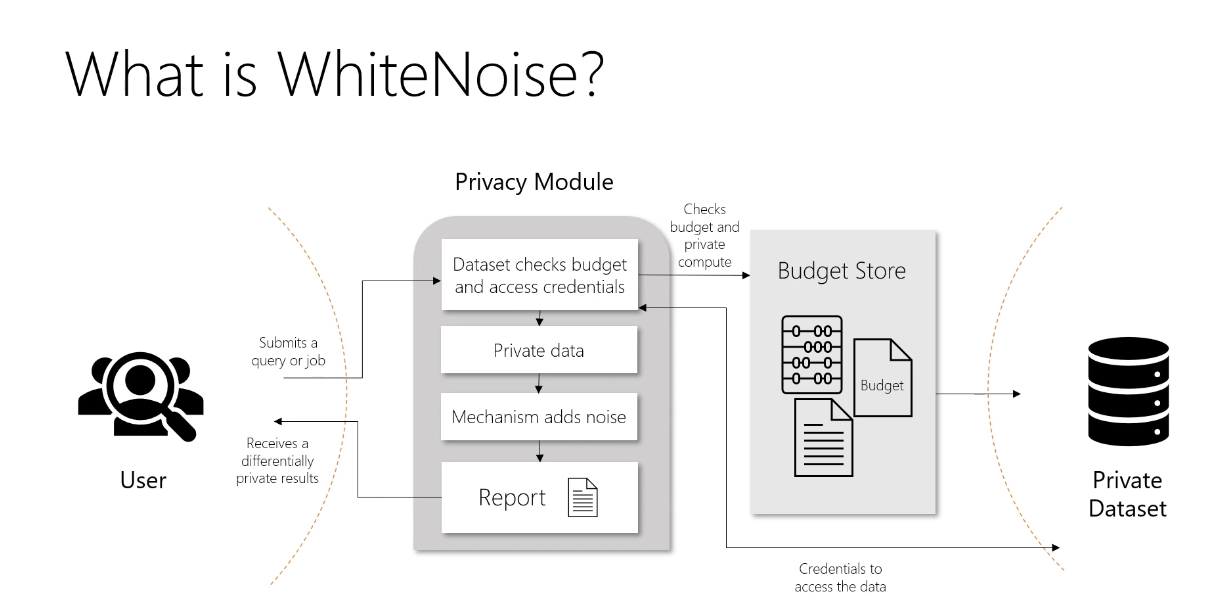
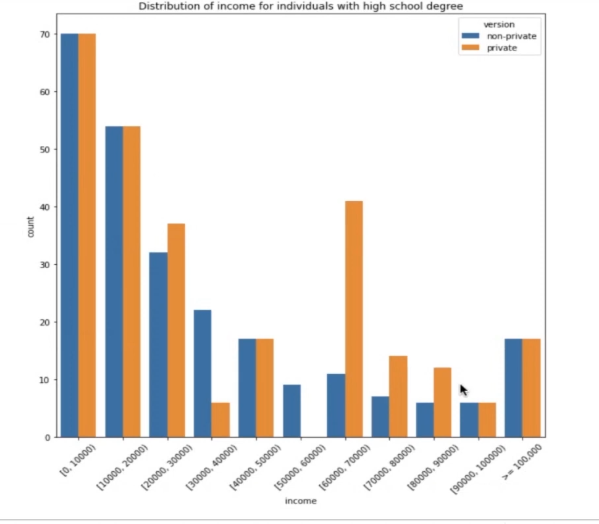
**Protecting Sensitive Data using Differential Privacy**

Sarah Bird

* Idea is to perform analytics over machine learning, while guaranteeing that you are not revealing any individual in the data set.
  + The model may memorise certain data points
* Example using income analysis on 500 people; ‘distribution of income for individuals with a high school degree’.
  + By using a simple statistics solver, and having access to some basic information of some data entries (e.g. there is at least one person in the data with ‘xyz’ details with an income of $95,000, there is ONLY on person with data with ‘abc’ details with an income of $31,000).
  + The stat solver correctly guessed the exact income of 54 people, correctly guessed the income within $2,000 for 87 people, and correctly guessed the income within $5,000 for 119 people.
  + By filling in information that’s an aggregate, and because a few records are known in the database, the stat solver is able to correctly guess the income for about a third of the individuals.
    - We would not know the specific third of individuals we have correctly identified unless the stat solver is run many times.
* Differential privacy algorithms take 2 steps to mask the contribution of individuals:
  + Add noise to the aggregate result.
    - This prevents us from getting the exact answer anymore, there will be some statistical noise to mask the contribution of individuals so that it cannot be traced back.
  + Track privacy budget.
    - If too many queries are made (each revealing a small amount of information) then it is possible for the original information to be recreated.
* “Are we literally flipping answers in the database on the ‘add-noise’ step?”. No, think of it as adding a small amount to the aggregates. For example, if the real answer to a query would be ‘60’, then a DP algorithm may return 61 or 65.



* In the future, hopeful that the data scientist cannot access the dataset directly; they will send a query to the privacy module which will return an answer with added noise.
* Running the same ‘statistical release and reconstruction attack’ using white noise returns ‘unsat’.
  + The program couldn’t find a database that satisfies the published reports so it has no guesses as to what an individual’s income may be.
* We still need to show that the data is useful! How useful is noisy data? So how can this noisy/private data still be used for the income analysis?



X variable is income, Y variable is the count of individuals within each income bracket.

* + In some cases, it’s quite close; the data is very usable, the noise is quite small.
  + In other cases, the noise is quite significant. This is for the smaller values because the relative noise being added could be a lot more.
  + Usability depends on how large the data set is and how much noise needs to be added (which will depend on the privacy needed).
* “Is there a slider between how much privacy you want vs how much noise is added?”. Yes, and how much budget you want to spend on a particular query; if you wanted to have one query and spend your entire budget on it, you would get a very accurate result vs many different queries that only use a small amount of the privacy budget. So the accuracy is determined by the privacy budget, but also the scale; a very large data set with have smaller variance between private and non-private