

Day 2: Variant detection pipeline

Loose ends; why code vs results

- ▶ No one asked me this why?

Different backup policies

- ▶ results files are
 - ▶ usually huge so can not backup very often
 - ▶ usually (should be) regenerable from the files in code and the raw data
- ▶ stuff in code is usually very precious and is being modified so basically what to back it up or snapshot it as often as possible.
 - ▶ code + data is in a sense a backup of results.
 - ▶ and code is usually small so can afford to back it up
- ▶ By separating the two make it possible to have different backup policies for each.

VC; snapshotting

Everyone do

```
cd ~/Day45
```

```
cat >.gitignore
```

```
results
```

```
^D #This means type a control-D
```

```
git init
```

```
git add .
```

```
git commit -m "initial check"
```