What computers can run Unix?

- Apple OS X Macs
- Google’s Android phones
- Wireless internet routers
- Airplane entertainment systems
The Terminal Window

the shell, the prompt, the command line
The Terminal Window

Make it comfortable to work in:

- Resize the window
- Change your font size
- Open multiple terminal windows
Obtain a cheat sheet

google “unix commands”

Basic UNIX commands

Note: not all of these are actually part of UNIX itself, and you may not find them on all UNIX machines. But they can all be used on turing in essentially the same way, by typing the command and hitting return. Note that some of these commands are different on non-Solaris machines - see SunOS differences.

If you've made a typo, the easiest thing to do is hit CTRL-u to cancel the whole line. But you can also edit the command line (see the guide to More UNIX).

UNIX is case-sensitive.

Files

- `ls` --- lists your files
  - `ls -l` --- lists your files in 'long format', which contains lots of useful information, e.g. the exact size of the file, who owns the file and who has the right to look at it, and when it was last modified.
  - `ls -a` --- lists all files, including the ones whose filenames begin in a dot, which you do not always want to see.

There are many more options, for example to list files by size, by date, recursively etc.

- `more filename` --- shows the first part of a file, just as much as will fit on one screen. Just hit the space bar to see more or q to quit. You can use /pattern to search for a pattern.

- `emacs filename` --- is an editor that lets you create and edit a file. See the emacs page.

- `mv filename1 filename2` --- moves a file (i.e. gives it a different name, or moves it into a different directory (see below)
- `cp filename1 filename2` --- copies a file

- `rm filename` --- removes a file. It is wise to use the option rm -i, which will ask you for confirmation before actually deleting anything.

You can make this your default by making an alias in your .cshrc file.

- `diff filename1 filename2` --- compares files, and shows where they differ
- `wc filename` --- tells you how many lines, words, and characters there are in a file

- `chmod options filename` --- lets you change the read, write, and execute permissions on your files. The default is that only you can look at them and change them, but you may sometimes want to change these permissions. For example, `chmod o+r filename` will make the file readable for everyone, and `chmod o=r filename` will make it unreadable for others again. Note that for someone to be able to actually look at the file the directories it is in need to be at least executable. See help protection for more details.

File Compression

- `gzip filename` --- compresses files, so that they take up much less space. Usually text files compress to about half their original size, but it depends very much on the size of the file and the nature of the content. There are other tools for this purpose, too (e.g. compress), but gzip usually gives the highest compression rate. Gzip produces files with the ending 'gz' appended to the original filename.

- `gunzip filename` --- uncompress file compressed by gzip.
In UNIX everything is a file organized in a hierarchy
Create a series of directories

```bash
% mkdir shell
% cd shell
% ls
research
% cd research/
% ls
seq
% cd seq/
% ls
radtags
% cd radtags/
% ls
```

```bash
total 8
drwxrwxr-x 2 ubuntu ubuntu 4096 2012-03-06 23:08 ..
drwxrwxr-x 3 ubuntu ubuntu 4096 2012-03-06 23:08 .
% pwd
/home/ubuntu/shell/research/seq/radtags
% ls -la
```
Paths, cont

This shell view of the nested directories shell, research, seq, and radtags..... is equivalent to this GUI view of the same directories.

And the **radtags** directory is uniquely identified by its path: `/home/ubuntu/shell/research/seq/radtags`
Absolute and relative paths

How do I get to the Hotel Zlaty Andel?
Absolute and relative paths

How do I get to the Hotel Zlaty Andel?
Absolute Path

/home/catchen/working
Relative Path?

```
/   
 |   
| home | usr | bin |
|      |     |     |
|      | cresko | catchen | lib | bin |
|      | working |             |
```

```bash
import os

# Define the base path
base_path = '/home/usr/bin...

# Define the working and research directories
working_dir = 'working
research_dir = 'research

# Calculate the relative path
relative_path = os.path.relpath(working_dir, base_path)
relative_path = os.path.relpath(research_dir, base_path)
```

The relative path from the base path to the working directory is `working`.

The relative path from the base path to the research directory is `research`.
Special files -- ‘dot’
Special files -- ‘dot dot’
Relative Path

```
/           
\root       
\home       
\usr        
\bin        
\cresko     
\catchen    
\lib        
\bin        
\working    
\research   
```

`../working`
Absolute Path: /home/catchen/working/foo
Relative Path: ../working/foo
Absolute Path: /home/catchen/working/foo
Absolute Path: /home/catchen/research/foo
Relative Path: ../../../working/foo
Relative Path: ../../foo
Absolute and relative paths

Special Files

dot
dot dot

% ls .
% ls ..
% ls ../../../
Binary programs - ls, cp, mkdir, etc.

% ls /bin
Relative and absolute paths

A shortcut to your ‘home’, tilde:

~

Moving through the filesystem:

cd

Knowing where you are:

pwd

% ls ~/
% cd ~/
% cd
% pwd
Relative and absolute paths

/home/tgac/shell/research/seq/radtags

% ls .
% ls ..
% ls ../../../

% cd ~/
% cd shell/research
% pwd
Are you typing? You’re doing it wrong.

Tab-completion:

- Tab once to complete uniquely
- Tab twice to see all possible completions

Up-arrow:

- Previous commands can be found by pressing “up-arrow”

‘history’

```
% cd /etc
% ls c <tab>
% pwd
% ls c <tab><tab>
```
<table>
<thead>
<tr>
<th></th>
<th>ls -l</th>
<th>ls -la</th>
<th>ls -lh</th>
</tr>
</thead>
<tbody>
<tr>
<td>Provides</td>
<td>provides a <em>long</em> listing</td>
<td>includes <em>all</em> files, even hidden files</td>
<td>displays file sizes in <em>human</em> readable numbers</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Four ways to view a text file

<table>
<thead>
<tr>
<th>more</th>
<th>head</th>
<th>tail</th>
<th>cat</th>
</tr>
</thead>
<tbody>
<tr>
<td>view a text file one</td>
<td>view the top 15 lines of a file</td>
<td>view the last 15 lines of a file</td>
<td>spit the whole file at once</td>
</tr>
<tr>
<td>screen full at a time</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>space-bar: scroll</td>
<td>-n num controls the number of</td>
<td>-n num controls the number of</td>
<td></td>
</tr>
<tr>
<td>q: quit</td>
<td>lines</td>
<td>lines</td>
<td></td>
</tr>
</tbody>
</table>
Explore the file hierarchy

1. Move to the directory /etc
   • What is the first line of the file ‘hosts’ in the directory /etc?
   • What is the relative file path to get to /var/log from here?
   • What is the absolute path?

2. Move to the directory /var/log/
   • What is the contents on line 73 of the dmesg file?
   • Without changing directories, what is the second line of the cpuinfo file in the proc directory?
     • What is the command to read this file with a relative path?
     • An absolute path?

3. Move back to the root, what directories do you see?

4. Move back home, what are three ways to move home from the root?
Copy example files

Return to the directory in your home called ‘shell’.

TSV file:
~/workshop_data/unix/batch_1.genotypes_1.loc.gz

FASTQ file:
~/workshop_data/unix/s_1_sequence.txt.gz

Tar Archive:
~/workshop_data/unix/samples.tar.gz
What is a tar archive?

tar = tape archive
Compress / Decompress

gzip / gunzip

batch_1.genotypes_1.loc.gz

s_1_sequence.txt.gz

Gzipped Tar archive

tar xvfz

samples.tar.gz

Tar archive

tar xvf

samples.tar
Sequencing on Illumina’s Flow cell
Sequencing on Illumina’s Flow cell, ctd.
Sequencing on Illumina’s Flow cell, ctd.

Phred Quality Score

$$Q = -10 \log_{10} p$$

**Phred Quality Score**

<table>
<thead>
<tr>
<th>Phred Quality Score</th>
<th>Probability of incorrect base call</th>
<th>Base call accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>1 in 10</td>
<td>90%</td>
</tr>
<tr>
<td>20</td>
<td>1 in 100</td>
<td>99%</td>
</tr>
<tr>
<td>30</td>
<td>1 in 1000</td>
<td>99.9%</td>
</tr>
<tr>
<td>40</td>
<td>1 in 10000</td>
<td>99.99%</td>
</tr>
<tr>
<td>50</td>
<td>1 in 100000</td>
<td>99.999%</td>
</tr>
</tbody>
</table>
The FASTQ File Format

FASTA

>chromosome7
TTTGTCTGCAGGGGGACACGTCAAAGTCAAACGCAGGCAAGTTTGTGTTTATGTCCAGTGGATCTTTTGATTTT
ACATACTGCAGGGTCAGGAGGATTATCTCTCTCTGCAAAGGTAAACGCTGCTGTAACCCTTGGTGTCTCTCATCCTTTTT
CCTAAGCAGGGGTCTGTCTGTCAGGCTCTGACAAGACATATGCAGGGCTCAATTGAGATAATTGCTCAATATA

FASTQ

@Sequence_137
TTTGTCTGCAGGGGGACACGTCAAAGTCAAACGCAGGCAAGTTTGTGTTTATGTCCAGTGGATCTTTTGATTTT
+Sequence_137
<?@DDDDDFHHFBB@G@ICFHGGHGBAGCDHBEAHACHI=CH.=7CAHHADDCDBCC66(6>@C>5@CACA

@HWI-ST0747:162:C03AJACXX:3:1108:19763:106771 1:N:0:
TTTGTCTGCAGGGGGACACGTCAAAGTCAAACGCAGGCAAGTTTGTGTTTATGTCCAGTGGATCTTTTGATTTT
+ #DFFHHFBB@G@ICFHGGHGBAGCDHBEAHACHI=CH.=7CAHHADDCDBCC66(6>@C>5@CACA
# ASCII Code

<table>
<thead>
<tr>
<th>Code</th>
<th>Character</th>
<th>Code</th>
<th>Character</th>
<th>Code</th>
<th>Character</th>
<th>Code</th>
<th>Character</th>
<th>Code</th>
<th>Character</th>
<th>Code</th>
<th>Character</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>&lt;NUL&gt;</td>
<td>64</td>
<td>@</td>
<td>96</td>
<td>`</td>
<td>128</td>
<td>Å</td>
<td>160</td>
<td>†</td>
<td>192</td>
<td>ć</td>
</tr>
<tr>
<td>1</td>
<td>&lt;SOH&gt;</td>
<td>65</td>
<td>A</td>
<td>97</td>
<td>a</td>
<td>129</td>
<td>Å</td>
<td>161</td>
<td>°</td>
<td>193</td>
<td>i</td>
</tr>
<tr>
<td>2</td>
<td>&lt;STX&gt;</td>
<td>66</td>
<td>B</td>
<td>98</td>
<td>b</td>
<td>130</td>
<td>Ç</td>
<td>162</td>
<td>ç</td>
<td>194</td>
<td>–</td>
</tr>
<tr>
<td>3</td>
<td>&lt;ETX&gt;</td>
<td>67</td>
<td>C</td>
<td>99</td>
<td>c</td>
<td>131</td>
<td>É</td>
<td>163</td>
<td>£</td>
<td>195</td>
<td>√</td>
</tr>
<tr>
<td>4</td>
<td>&lt;EOT&gt;</td>
<td>68</td>
<td>D</td>
<td>100</td>
<td>d</td>
<td>132</td>
<td>Ñ</td>
<td>164</td>
<td>§</td>
<td>196</td>
<td>f</td>
</tr>
<tr>
<td>5</td>
<td>&lt;ENQ&gt;</td>
<td>69</td>
<td>E</td>
<td>101</td>
<td>e</td>
<td>133</td>
<td>Ö</td>
<td>165</td>
<td>•</td>
<td>197</td>
<td>≈</td>
</tr>
<tr>
<td>6</td>
<td>&lt;ACK&gt;</td>
<td>70</td>
<td>F</td>
<td>102</td>
<td>f</td>
<td>134</td>
<td>Ú</td>
<td>166</td>
<td>‑</td>
<td>198</td>
<td>Δ</td>
</tr>
<tr>
<td>7</td>
<td>&lt;BEL&gt;</td>
<td>71</td>
<td>G</td>
<td>103</td>
<td>g</td>
<td>135</td>
<td>á</td>
<td>167</td>
<td>B</td>
<td>199</td>
<td>«</td>
</tr>
<tr>
<td>8</td>
<td>&lt;BS&gt;</td>
<td>72</td>
<td>H</td>
<td>104</td>
<td>h</td>
<td>136</td>
<td>à</td>
<td>168</td>
<td>@</td>
<td>200</td>
<td>»</td>
</tr>
<tr>
<td>9</td>
<td>&lt;TAB&gt;</td>
<td>73</td>
<td>I</td>
<td>105</td>
<td>i</td>
<td>137</td>
<td>â</td>
<td>169</td>
<td>©</td>
<td>201</td>
<td>…</td>
</tr>
<tr>
<td>10</td>
<td>&lt;LF&gt;</td>
<td>74</td>
<td>J</td>
<td>106</td>
<td>j</td>
<td>138</td>
<td>å</td>
<td>170</td>
<td>′</td>
<td>202</td>
<td>2</td>
</tr>
<tr>
<td>11</td>
<td>&lt;VT&gt;</td>
<td>75</td>
<td>K</td>
<td>107</td>
<td>k</td>
<td>139</td>
<td>ä</td>
<td>171</td>
<td>′</td>
<td>203</td>
<td>À</td>
</tr>
<tr>
<td>12</td>
<td>&lt;FF&gt;</td>
<td>76</td>
<td>L</td>
<td>108</td>
<td>l</td>
<td>140</td>
<td>Ñ</td>
<td>172</td>
<td>″</td>
<td>204</td>
<td>Â</td>
</tr>
<tr>
<td>13</td>
<td>&lt;CR&gt;</td>
<td>77</td>
<td>M</td>
<td>109</td>
<td>m</td>
<td>141</td>
<td>ç</td>
<td>173</td>
<td>≠</td>
<td>205</td>
<td>Ō</td>
</tr>
<tr>
<td>14</td>
<td>&lt;SO&gt;</td>
<td>78</td>
<td>N</td>
<td>110</td>
<td>n</td>
<td>142</td>
<td>é</td>
<td>174</td>
<td>Æ</td>
<td>206</td>
<td>Ê</td>
</tr>
<tr>
<td>15</td>
<td>&lt;SI&gt;</td>
<td>79</td>
<td>O</td>
<td>111</td>
<td>o</td>
<td>143</td>
<td>è</td>
<td>175</td>
<td>Ø</td>
<td>207</td>
<td>œ</td>
</tr>
<tr>
<td>16</td>
<td>&lt;DLE&gt;</td>
<td>80</td>
<td>P</td>
<td>112</td>
<td>p</td>
<td>144</td>
<td>è</td>
<td>176</td>
<td>∞</td>
<td>208</td>
<td>ő</td>
</tr>
<tr>
<td>17</td>
<td>&lt;DC1&gt;</td>
<td>81</td>
<td>Q</td>
<td>113</td>
<td>q</td>
<td>145</td>
<td>ê</td>
<td>177</td>
<td>±</td>
<td>209</td>
<td>−</td>
</tr>
<tr>
<td>18</td>
<td>&lt;DC2&gt;</td>
<td>82</td>
<td>R</td>
<td>114</td>
<td>r</td>
<td>146</td>
<td>í</td>
<td>178</td>
<td>≤</td>
<td>210</td>
<td>&quot;</td>
</tr>
<tr>
<td>19</td>
<td>&lt;DC3&gt;</td>
<td>83</td>
<td>S</td>
<td>115</td>
<td>s</td>
<td>147</td>
<td>i</td>
<td>179</td>
<td>≥</td>
<td>211</td>
<td>&quot;</td>
</tr>
<tr>
<td>20</td>
<td>&lt;DC4&gt;</td>
<td>84</td>
<td>T</td>
<td>116</td>
<td>t</td>
<td>148</td>
<td>i</td>
<td>180</td>
<td>¥</td>
<td>212</td>
<td>ʿ</td>
</tr>
<tr>
<td>21</td>
<td>&lt;NAK&gt;</td>
<td>85</td>
<td>U</td>
<td>117</td>
<td>u</td>
<td>149</td>
<td>ī</td>
<td>181</td>
<td>μ</td>
<td>213</td>
<td>ʿ</td>
</tr>
<tr>
<td>22</td>
<td>&lt;SYN&gt;</td>
<td>86</td>
<td>V</td>
<td>118</td>
<td>v</td>
<td>150</td>
<td>ņ</td>
<td>182</td>
<td>ḥ</td>
<td>214</td>
<td>ⅈ</td>
</tr>
<tr>
<td>23</td>
<td>&lt;ETB&gt;</td>
<td>87</td>
<td>W</td>
<td>119</td>
<td>w</td>
<td>151</td>
<td>ó</td>
<td>183</td>
<td>Σ</td>
<td>215</td>
<td>ᵗ</td>
</tr>
<tr>
<td>24</td>
<td>&lt;CAN&gt;</td>
<td>88</td>
<td>X</td>
<td>120</td>
<td>x</td>
<td>152</td>
<td>ő</td>
<td>184</td>
<td>Π</td>
<td>216</td>
<td>̃</td>
</tr>
<tr>
<td>25</td>
<td>&lt;EM&gt;</td>
<td>89</td>
<td>Y</td>
<td>121</td>
<td>y</td>
<td>153</td>
<td>ó</td>
<td>185</td>
<td>n</td>
<td>217</td>
<td>̇</td>
</tr>
<tr>
<td>26</td>
<td>&lt;SUB&gt;</td>
<td>90</td>
<td>Z</td>
<td>122</td>
<td>z</td>
<td>154</td>
<td>ö</td>
<td>186</td>
<td>f</td>
<td>218</td>
<td>ⅉ</td>
</tr>
<tr>
<td>27</td>
<td>&lt;ESC&gt;</td>
<td>91</td>
<td>[</td>
<td>123</td>
<td>{</td>
<td>155</td>
<td>ɨ</td>
<td>187</td>
<td>a</td>
<td>219</td>
<td>ʾ</td>
</tr>
<tr>
<td>28</td>
<td>&lt;FS&gt;</td>
<td>92</td>
<td>\</td>
<td>124</td>
<td></td>
<td></td>
<td>156</td>
<td>ú</td>
<td>188</td>
<td>o</td>
<td>220</td>
</tr>
<tr>
<td>29</td>
<td>&lt;GS&gt;</td>
<td>93</td>
<td>]</td>
<td>125</td>
<td>}</td>
<td>157</td>
<td>ū</td>
<td>189</td>
<td>Ω</td>
<td>221</td>
<td>&gt;</td>
</tr>
<tr>
<td>30</td>
<td>&lt;RS&gt;</td>
<td>94</td>
<td>^</td>
<td>126</td>
<td>~</td>
<td>158</td>
<td>ü</td>
<td>190</td>
<td>æ</td>
<td>222</td>
<td>fi</td>
</tr>
<tr>
<td>31</td>
<td>&lt;US&gt;</td>
<td>95</td>
<td>_</td>
<td>127</td>
<td>&lt;DEL&gt;</td>
<td>159</td>
<td>ü</td>
<td>191</td>
<td>ø</td>
<td>223</td>
<td>fl</td>
</tr>
</tbody>
</table>

8 bits = $2^8$ combinations = 256

<table>
<thead>
<tr>
<th>Code</th>
<th>Character</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

0x$2^7$+1x$2^6$+1x$2^5$+1x$2^4$+1x$2^3$+0x$2^2$+0x$2^1$+1x$2^0$ = 121 = y

1x$10^2$+2x$10^1$+1x$10^0$ = 121 = y
The FASTQ File Format, ctd

@HWI-ST0747:162:C03AJACXX:3:1108:19763:106771 1:N:0:
TTTGTCGTGGACGGGGACACGTCAAGTCAAACGCAGGCAAGTTTGTTATGTCCAGTGGATCTTTTGATTTT+

<@DDDDDFHHHFB@GGIACFHGGHBHGCDHBEAHACHI=@CH.=7ACAHHADECDBCC66(6>@C>5@CACCA

Quality Scores

<table>
<thead>
<tr>
<th>ASCII values</th>
<th>S - Sanger Phred+33, raw reads typically (0, 40)</th>
<th>X - Solexa Solexa+64, raw reads typically (-5, 40)</th>
<th>I - Illumina 1.3+ Phred+64, raw reads typically (0, 40)</th>
<th>J - Illumina 1.5+ Phred+64, raw reads typically (3, 40)</th>
</tr>
</thead>
<tbody>
<tr>
<td>33 - 73</td>
<td>0 - 40</td>
<td>-5 - 40</td>
<td>0 - 40</td>
<td>3 - 40</td>
</tr>
</tbody>
</table>

with 0=unused, 1=unused, 2=Read Segment Quality Control Indicator (bold)
(Note: See discussion above).

L - Illumina 1.8+ Phred+33, raw reads typically (0, 41)

ASCII values 33 - 73 = 0 - 40

‘F’ = 70 70 - 33 = 37

http://en.wikipedia.org/wiki/FASTQ_format
The FASTQ File Format, ctd

@HWI-ST0747:162:C03AJACXX:3:1108:19763:106771 1:N:0:
TTTGTCTGCAAGGGGACACGTCAAAGTCAAACGCAGGCAGTATTATGTGGATTGATTTTGATTTT
+
<?@DDDDDFHHFBB@GGIACFHGGHBGGCDHBEAHACHI=@CH.=7ACAHHADECDBCC66(6>@C>5@CACCA

70 - 33 = 37

<table>
<thead>
<tr>
<th>Phred Quality Score</th>
<th>Probability of incorrect base call</th>
<th>Base call accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>1 in 10</td>
<td>90%</td>
</tr>
<tr>
<td>20</td>
<td>1 in 100</td>
<td>99%</td>
</tr>
<tr>
<td>30</td>
<td>1 in 1000</td>
<td>99.9%</td>
</tr>
<tr>
<td>40</td>
<td>1 in 10000</td>
<td>99.99%</td>
</tr>
<tr>
<td>50</td>
<td>1 in 100000</td>
<td>99.999%</td>
</tr>
</tbody>
</table>

http://en.wikipedia.org/wiki/Phred_quality_score
Count raw reads:

```
wc -l s_1_sequence.txt
```

```
grep "@" s_1_sequence.txt
grep -c "@" s_1_sequence.txt
```

```
grep -v "@" s_1_sequence.txt
grep -v -c "@" s_1_sequence.txt
```

Count reads with barcode:

```
grep -c "^CGATA" s_1_sequence.txt
```
Special Files

STDOUT, STDERR

The Shell’s Killer App: Pipes

Leci n’est pas une pipe.
The Shell’s Killer App: **Pipes, ctd.**

STDIN → **Program I** → STDOUT
The Shell’s Killer App: **Pipes**, ctd.

So what is the purpose of the program **cat**?
cut

`cut -f 10 batch_1.genotypes_1.loc`

cut, capture the output

`cut -f 1-10 batch_1.genotypes_1.loc > genos`

cut, pipe the output to grep

`cut -f 2 batch_1.genotypes_1.loc | grep -c "nnxnp"`

`cut -f 1-10,15,17 batch_1.genotypes_1.loc | grep "nnxnp" > genos2`

Examine a marker, translating the output

`cat batch_1.genotypes_1.loc | tr " " "," | grep "^96053"`
Decompress the file

1. Count the number of raw reads (250,000)
2. Count the number of reads with barcode CGATA (19,501)
3. Capture all FASTQ records for ACCAT into a file called sample_01.fq (you should get 18352 records, 73408 lines)
4. Determine the count of all barcodes in the file

- 286 CTAGT
- 7900 TCAGA
- 10659 ACTGC
- 10931 TGACC
- 11536 GAGAT
- 11871 CTGAA
- 14409 CGGCG
- 14508 TGGTT
- 18226 GAAGC
- 18352 ACCAT
- 18375 TCGAG
- 19501 CGATA
- 23012 AATTT
- 26336 GCATT
- 31136 CTAGG

1. Use `head` when building a command, `cat` once the command is working
2. Look at the `-n` option for the `head` command, the `-l` option for `wc`
3. The “^” character means “must occur at beginning of line” in a `grep` search
4. Look at the `grep` options: `-c`, `-v`, `-A`, `-B`
5. Read the man pages for `sort` and `uniq` to learn how to combine them
<table>
<thead>
<tr>
<th># Batch ID</th>
<th>Locus ID</th>
<th>Pop 1 ID</th>
<th>Pop 2 ID</th>
<th>Chr</th>
<th>BP</th>
<th>Column</th>
<th>Overall Pi</th>
<th>Fst</th>
<th>Fisher's P</th>
<th>Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>7894</td>
<td>2</td>
<td>3</td>
<td>groupI</td>
<td>11832</td>
<td>19</td>
<td>0.428182</td>
<td>-0.0076252913</td>
<td>0.191294</td>
<td>0.687192</td>
</tr>
<tr>
<td>2</td>
<td>7896</td>
<td>2</td>
<td>3</td>
<td>groupI</td>
<td>11900</td>
<td>83</td>
<td>0.328622</td>
<td>0.1775694587</td>
<td>4.35747e-08</td>
<td>5.44667</td>
</tr>
<tr>
<td>2</td>
<td>9611</td>
<td>2</td>
<td>3</td>
<td>groupI</td>
<td>49756</td>
<td>48</td>
<td>0.090426</td>
<td>-0.1127451906</td>
<td>0.00109115</td>
<td>0.0898072</td>
</tr>
<tr>
<td>2</td>
<td>9611</td>
<td>2</td>
<td>3</td>
<td>groupI</td>
<td>49765</td>
<td>57</td>
<td>0.0132887</td>
<td>-0.1522407447</td>
<td>0.518395</td>
<td>0.493113</td>
</tr>
<tr>
<td>2</td>
<td>9611</td>
<td>2</td>
<td>3</td>
<td>groupI</td>
<td>49766</td>
<td>58</td>
<td>0.0133776</td>
<td>-0.1563725438</td>
<td>0.520033</td>
<td>0.501401</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>7083</td>
<td>2</td>
<td>3</td>
<td>groupIX</td>
<td>20172984</td>
<td>57</td>
<td>0.0480227</td>
<td>-0.1249752728</td>
<td>0.0230205</td>
<td>0.152263</td>
</tr>
<tr>
<td>2</td>
<td>7083</td>
<td>2</td>
<td>3</td>
<td>groupIX</td>
<td>20173004</td>
<td>77</td>
<td>0.0181808</td>
<td>-0.1405892833</td>
<td>0.27288</td>
<td>0.355556</td>
</tr>
<tr>
<td>2</td>
<td>7083</td>
<td>2</td>
<td>3</td>
<td>groupIX</td>
<td>20173016</td>
<td>89</td>
<td>0.0424072</td>
<td>-0.1243625038</td>
<td>0.0442229</td>
<td>0.17037</td>
</tr>
<tr>
<td>2</td>
<td>7083</td>
<td>2</td>
<td>3</td>
<td>groupIX</td>
<td>20173018</td>
<td>91</td>
<td>0.148072</td>
<td>0.2305903127</td>
<td>2.24832e-07</td>
<td>13.2936</td>
</tr>
<tr>
<td>2</td>
<td>7084</td>
<td>2</td>
<td>3</td>
<td>groupIX</td>
<td>20172960</td>
<td>29</td>
<td>0.388239</td>
<td>-0.0102170169</td>
<td>0.307185</td>
<td>0.758001</td>
</tr>
<tr>
<td>2</td>
<td>7084</td>
<td>2</td>
<td>3</td>
<td>groupIX</td>
<td>20172995</td>
<td>64</td>
<td>0.368259</td>
<td>-0.0130602425</td>
<td>0.233889</td>
<td>0.70084</td>
</tr>
<tr>
<td>2</td>
<td>7084</td>
<td>2</td>
<td>3</td>
<td>groupIX</td>
<td>20173013</td>
<td>82</td>
<td>0.0121578</td>
<td>-0.1359057575</td>
<td>0.51309</td>
<td>0.46472</td>
</tr>
</tbody>
</table>