

# Reproducible Research Checklist

Adapted from:

[https://github.com/DataScienceSpecialization/courses/blob/master/05\\_ReproducibleResearch/Checklist/Reproducible%20Research%20Checklist.pdf](https://github.com/DataScienceSpecialization/courses/blob/master/05_ReproducibleResearch/Checklist/Reproducible%20Research%20Checklist.pdf)

<b>Think about the entire pipeline:</b> are all the pieces reproducible?	
<b>Is your cleaning/analysis process automated?</b> – guarantees reproducibility <ul style="list-style-type: none"><li>• <u>Are you doing things “by hand”?</u> editing tables/figures; splitting/reformatting data</li><li>• <u>Does your software support log files or scripts?</u></li><li>• <u>If no, do you have a detailed description of your process?</u></li></ul>	
<b>Are you using version control?</b> <ul style="list-style-type: none"><li>• Saving copies</li><li>• git/bitbucket/sourceforge</li></ul>	
<b>Are you keeping track of your software?</b> <ul style="list-style-type: none"><li>• Computer architecture;</li><li>• OS/Software/tool/add ons (libraries/packages)/external databases</li><li>• version numbers for everything (when available)</li></ul>	
<b>Are you saving the right files?</b> <ul style="list-style-type: none"><li>• Raw data</li><li>• Processing Code</li><li>• Processed data</li><li>• Results (tables and graphs) don't need to be saved: they can be regenerated because your analysis is reproducible!</li></ul>	
<b>Are your reports human and machine readable?</b> <ul style="list-style-type: none"><li>• Scripts/code</li><li>• Narrative</li></ul>	