

# The project title goes here

YOUR NAME<sup>1</sup>, COAUTHOR ONE<sup>2</sup>, AND COAUTHOR TWO<sup>1,3</sup>

<sup>1</sup>The first affiliation in the list.; [your@email.address](mailto:your@email.address)

<sup>2</sup>Another affiliation can be put here.

<sup>3</sup>And generally as many affiliations as you like.

*Received YYYY MM DD; accepted YYYY MM DD; published YYYY MM DD*

Welcome to Maneage (*Managing data lineage*) and reproducible papers/projects, for a review of the basics of this system, please see Akhlaghi et al. (2020). You are now ready to configure Maneage and implement your own research in this framework. Maneage contains almost all the elements that you will need in a research project, and adding any missing parts is very easy once you become familiar with it. For example it already has steps to downloading of raw data and necessary software (while verifying them with their checksums), building the software, and processing the data with the software in a highly-controlled environment. But Maneage is not just for the analysis of your project, you will also write your paper in it (by replacing this text in `paper.tex`): including this abstract, figures and bibliography. If you design your project with Maneage's infra-structure, don't forget to add a notice and clearly let the readers know that your work is reproducible, we should spread the word and show the world how useful reproducible research is for the sciences, also don't forget to cite and acknowledge it so we can continue developing it. This PDF was made with Maneage, commit db2bd88.

*Keywords:* Add some keywords for your research here.

*Reproducible paper:* All quantitative results (numbers and plots) in this paper are exactly reproducible with Maneage (<https://maneage.org>).

## 1. CONGRATULATIONS!

Congratulations on running the raw template project! You can now follow the “Customization checklist” in the `README-hacking.md` file, customize this template and start your exciting research project over it. You can always merge Maneage back into your project to improve its infra-structure and leaving your own project intact. If you haven't already read Akhlaghi et al. (2020), please do so before continuing, it isn't long (just 7 pages).

While you are writing your paper, just don't forget to *not* use numbers or fixed strings (for example database urls like <https://fits.gsfc.nasa.gov/samples/WFPC2ASSNu5780205bx.fits>) directly within your  $\text{\LaTeX}$  source. Put them in configuration files and after using them in the analysis, pass them into the  $\text{\LaTeX}$  source through macros in the same `subMakefile` that used them. For some already published examples, please see Akhlaghi et al. (2020)<sup>1</sup>, Infante-Sainz et al. (2020)<sup>2</sup> and Akhlaghi (2019)<sup>3</sup>. Working in this way, will let you focus clearly on your science and not have to worry about fixing this or that number/name in the text.

Once your project is ready for publication, there is also a “Publication checklist” in `README-hacking.md` that will guide you in the steps to do for making your project as FAIR as possible (Findable, Accessible, Interoperable, and Reusable).

The default  $\text{\LaTeX}$  structure within Maneage also has two  $\text{\LaTeX}$  macros for easy marking of text within your document as *new* and

*notes*. For example, **this text has been marked as new.** [While this one is marked as *tonote*.] Please try commenting the line that defines `highlightchanges` in `paper.tex` (by adding a `'%`' at the start of the line or simply deleting the line). You will then notice that the line that was marked as *new* will become black (totally blend in with the rest of the text) and the one marked *tonote* will not be in the final PDF. You can thus use `highlightchanges` to easily make copies of your research for existing coauthors (who are just interested in the new parts or notes) and new co-authors (who don't want to be distracted by these issues in their first time reading).

Figure 1 shows a simple plot as a demonstration of creating plots within  $\text{\LaTeX}$  (using the PGFPlots package). The minimum value in this distribution is 1.000, and 2500.000 is the maximum. Take a look into the  $\text{\LaTeX}$  source and you'll see these numbers are actually macros that were calculated from the same dataset (they will change if the dataset, or function that produced it, changes).

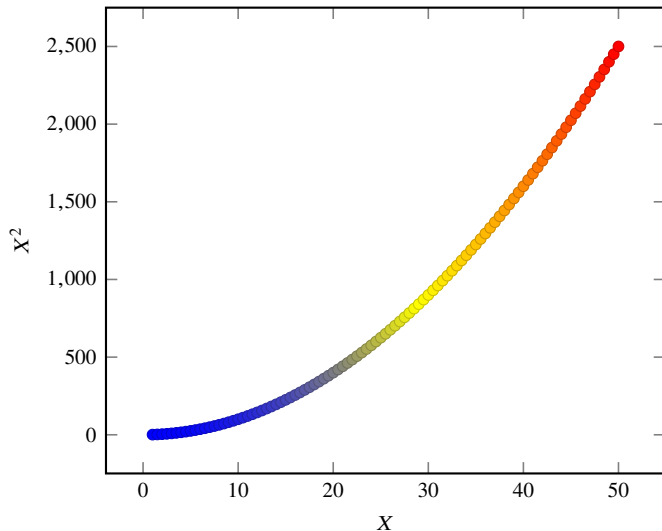
The individual PDF file of Figure 1 is available under the `tex/tikz/` directory of your build directory. You can use this PDF file in other contexts (for example in slides showing your progress or after publishing the work). If you want to directly use the PDF file in the figure without having to let TiKZ decide if it should be remade or not, you can also comment the `makepdf` macro at the top of this  $\text{\LaTeX}$  source file.

Figure 2 is another demonstration of showing images (datasets) using PGFPlots. It shows a small crop of an image from the Wide-Field Planetary Camera 2 (that was installed on the Hubble Space Telescope from 1993 to 2009). As another more realistic demonstration of reporting results with Maneage, here we report

<sup>1</sup> <https://gitlab.com/makhlaghi/maneage-paper>

<sup>2</sup> <https://gitlab.com/infantesainz/sdss-extended-psfs-paper>

<sup>3</sup> <https://gitlab.com/makhlaghi/iau-symposium-355>



**Figure 1:** A very basic  $X^2$  plot for demonstration.

that the mean pixel value in that image is 5.12 and the median is 0.86. The skewness in the histogram of Figure 2(b) explains this difference between the mean and median. The dataset is visualized here as a black and white image using the Convert-Type program of GNU Astronomy Utilities (Gnuastro). The histogram and basic statistics were generated with Gnuastro’s Statistics program.

PGFPlots<sup>4</sup> is a great tool to build the plots within  $\text{\LaTeX}$  and removes the necessity to add further dependencies, just to create the plots. There are high-level libraries like Matplotlib which also generate plots. However, the problem is that they require *many* dependencies, for example see Figure 1 of Alliez et al. (2020). Installing these dependencies from source, is not easy and will harm the reproducibility of your paper in the future.

Furthermore, since PGFPlots builds the plots within  $\text{\LaTeX}$ , it respects all the properties of your text (for example line width and fonts and etc). Therefore the final plot blends in your paper much more nicely. It also has a wonderful manual<sup>5</sup>.

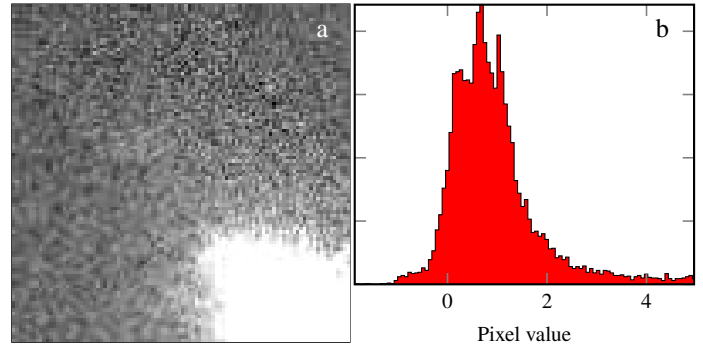
## 2. NOTICE AND CITATIONS

To encourage other scientists to publish similarly reproducible papers, please add a notice close to the start of your paper or in the end of the abstract clearly mentioning that your work is fully reproducible. One convention we have adopted until now is to put the Git checksum of the project as the last word of the abstract, for example see Akhlaghi (2019), Infante-Sainz et al. (2020) and Akhlaghi et al. (2020)

Finally, don’t forget to cite Akhlaghi et al. (2020) and acknowledge the funders mentioned below. Otherwise we won’t be able to continue working on Maneage. Also, just as another reminder, before publication, don’t forget to follow the “Publication checklist” of README-hacking.md.

<sup>4</sup> <https://ctan.org/pkg/pgfplots>

<sup>5</sup> <http://mirrors.ctan.org/graphics/pgf/contrib/pgfplots/doc/pgfplots.pdf>



**Figure 2:** (a) An example image of the Wide-Field Planetary Camera 2, on board the Hubble Space Telescope from 1993 to 2009. This is one of the sample images from the FITS standard webpage, kept as examples for this file format. (b) Histogram of pixel values in (a).

## 3. ACKNOWLEDGEMENTS

Please include the following two paragraphs in the Acknowledgement section of your paper. Maneage was developed in parallel with Gnuastro, so it benefited from the same grants. If you don’t use Gnuastro in your final/customized project, please remove it from the paragraph below, only mentioning the reproducible paper template.

This research was partly done using GNU Astronomy Utilities (Gnuastro, [ascl.net/1801.009](http://ascl.net/1801.009)), and the reproducible paper template db2bd88. Work on Gnuastro and the reproducible paper template has been funded by the Japanese Ministry of Education, Culture, Sports, Science, and Technology (MEXT) scholarship and its Grant-in-Aid for Scientific Research (21244012, 24253003), the European Research Council (ERC) advanced grant 339659-MUSICOS, European Union’s Horizon 2020 research and innovation programme under Marie Skłodowska-Curie grant agreement No 721463 to the SUNDIAL ITN, and from the Spanish Ministry of Economy and Competitiveness (MINECO) under grant number AYA2016-76219-P. The reproducible paper template was also supported by European Union’s Horizon 2020 (H2020) research and innovation programme via the RDA EU 4.0 project (ref. GA no. 777388).

## References

- Akhlaghi, M. and T. Ichikawa (2015). *ApJS*, 220, 1.
- Akhlaghi, M. (2019). *IAU Symposium* 335, [arXiv:1909.11230](https://arxiv.org/abs/1909.11230).
- Akhlaghi, M. et al. (2020). *arXiv e-prints*, [arXiv:2006.03018](https://arxiv.org/abs/2006.03018).
- Alliez, P. et al. (2020). *CiSE*, 22, 39–52.
- Infante-Sainz, R. et al. (2020). *MNRAS*, 491, 5317–5329.

## A. SOFTWARE ACKNOWLEDGEMENT

This research was done with the following free software programs and libraries: Bzip2 1.0.6, C compiler (gcc (GCC) 10.1.0), CFITSIO 3.47, CMake 3.17.0, cURL 7.65.3, Dash 0.5.10.2, Discoteq flock 0.2.3, File 5.36, Git 2.26.2, GNU Astronomy Utilities 0.11 (Akhlaghi and Ichikawa 2015), GNU AWK 5.0.1, GNU Bash 5.0.11, GNU Binutils 2.32, GNU Coreutils 8.31, GNU Diffutils 3.7, GNU Findutils 4.7.0, GNU gettext 0.20.2, GNU Grep

3.3, GNU Gzip 1.10, GNU Integer Set Library 0.18, GNU libiconv 1.16, GNU Libtool 2.4.6, GNU libunistring 0.9.10, GNU M4 1.4.18, GNU Make 4.3, GNU Multiple Precision Arithmetic Library 6.1.2, GNU Multiple Precision Complex library, GNU Multiple Precision Floating-Point Reliably 4.0.2, GNU NCURSES 6.1, GNU Readline 8.0, GNU Scientific Library 2.6, GNU Sed 4.7, GNU Tar 1.32, GNU Texinfo 6.6, GNU Wget 1.20.3, GNU Which 2.21, GPL Ghostscript 9.50, Libbsd 0.9.1, Libgit2 0.28.2, Libjpeg v9b, Libpng 1.6.37, Libtiff 4.0.10, Libxml2 2.9.9, Lzip 1.20, Metastore (forked) 1.1.2-23-fa9170b, OpenSSL 1.1.1a, PatchELF 0.10, Perl 5.30.0, pkg-config 0.29.2, Unzip 6.0, WCSLIB 6.4, XZ Utils 5.2.4, Zip 3.0 and Zlib 1.2.11. The  $\LaTeX$  source of the paper was compiled to make the PDF using the following packages: biber 2.14, biblatex 3.14, bitset 1.3, caption 55117 (revision), courier 35058 (revision), csquotes 5.2j, datetime 2.60, ec 1.0, etoolbox 2.5h, fancyhdr 3.10, fmtcount 3.07, fontaxes 1.0d, footmisc 5.5b, fp 2.1d, kastrup 15878 (revision), letltxmacro 1.6, logreq 1.0, newtx 1.627, pdfescape 1.15, pdftexcmds 0.31, pgf 3.1.5b, pgfplots 1.17, preprint 2011, setspace 6.7a, tex 3.14159265, texgyre 2.501, times 35058 (revision), titlesec 2.13, trimspaces 1.1, ttfonts 15878 (revision), ulem 53365 (revision), xcolor 2.12 and xkeyval 2.7a. We are very grateful to all their creators for freely providing this necessary infrastructure. This research (and many others) would not be possible without them.