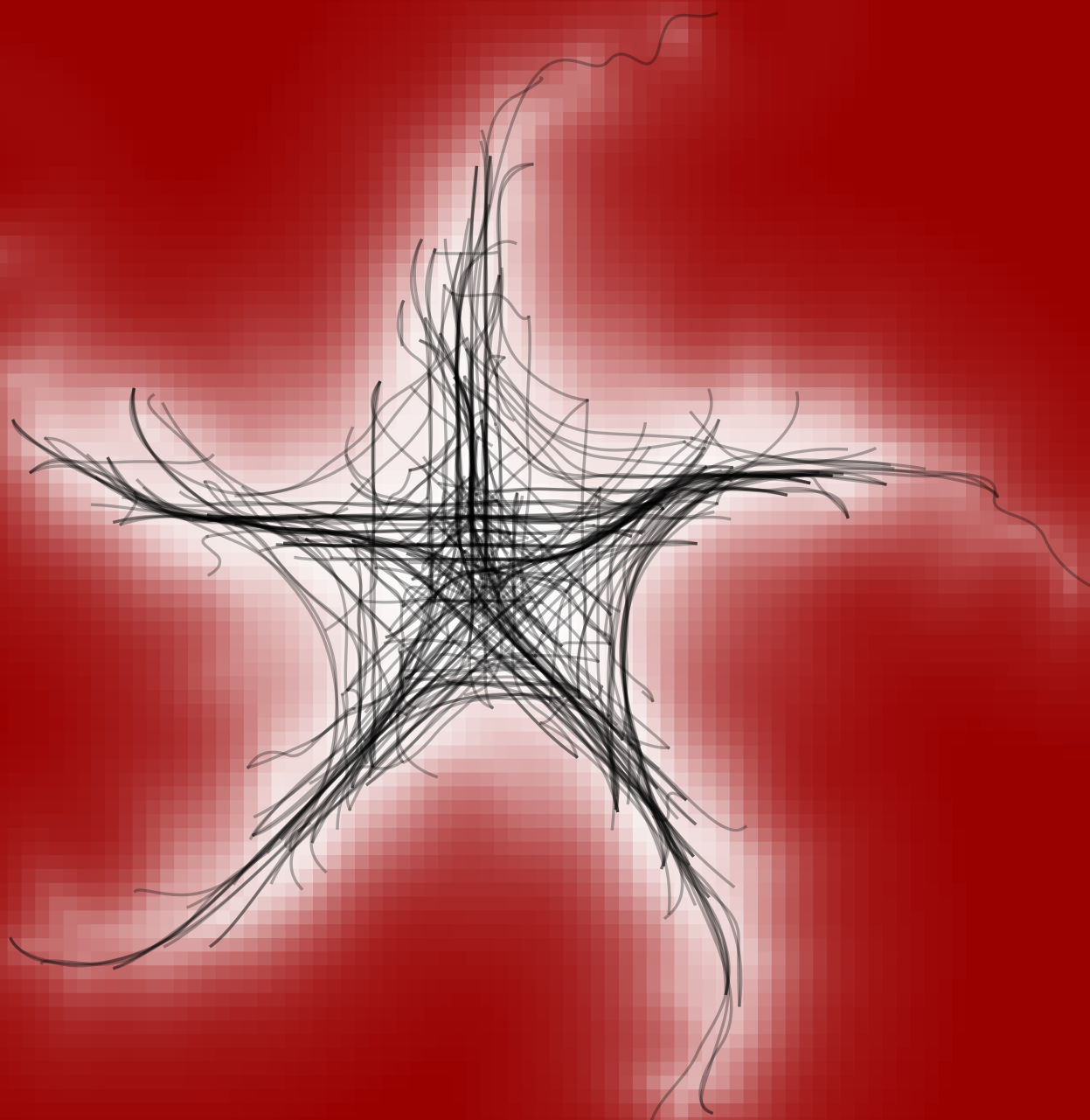


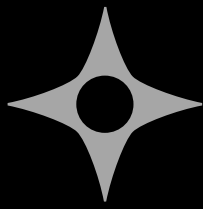


DTU COMPUTE
DEPARTMENT OF APPLIED MATHEMATICS AND COMPUTER SCIENCE

STATISTICS UNDER STOCHASTIC METRICS

CILIE W. FELDAGER
PH.D. THESIS





PH.D. THESIS
STATISTICS UNDER STOCHASTIC METRICS
BY CILIE W. FELDAGER

COGNITIVE SYSTEMS
DTU COMPUTE
TECHNICAL UNIVERSITY OF DENMARK

SUPERVISED BY
PROFESSOR SØREN HAUBERG
PROFESSOR LARS KAI HANSEN

CONTENTS

| | | |
|-----|-------------------------|----|
| 1 | Documentation of class | 1 |
| 1.1 | Main.tex | 1 |
| 1.2 | Preamble | 2 |
| 1.3 | Frontmatter | 3 |
| 1.4 | Backmatter | 3 |
| 2 | My writing process | 5 |
| 2.1 | The Braindump | 5 |
| 2.2 | The Outline | 5 |
| 2.3 | The Vomit Draft | 5 |
| 2.4 | Trees maps and theorems | 5 |
| 2.5 | The paragraph | 6 |
| 2.6 | Sanity checks | 7 |
| 3 | Random Examples | 9 |
| 3.1 | Short title for sec | 9 |
| 3.2 | Other Examples | 9 |
| 3.3 | Remarks and Theorems | 10 |
| 4 | Experiments | 15 |

APPENDIX

DOCUMENTATION OF CLASS

1

“ *That, if a straight line falling on two straight lines makes the interior angles on the same side less than two right angles, the two straight lines, if produced indefinitely, meet on that side on which the angles are less than two right angles.*

”

—Euclid

This class allows for typesetting a beautiful thesis with a few custom options. This chapter documents and displays the options in this class I implemented. Feel free to make it your own.

It builds on two other classes, Jesper’s and Dion’s which is build on ?? and uses the principles from `Trees`, `maps`, and `theorems` by Jean-Luc Dumont (see section 2.4). Ultimately this are build on `memoir`. This can be used for any type of document.

1.1 MAIN.TEX

For convenience, I introduced booleans that control if frontmatter, main matter or backmatter should be compiled. This is handy when having with numerous input files but working only in one.

1.1.1 Blackthesis

The `blackthesis` is an option parsed in the document class, eg `\documentclass[a4paper,11pt,twoside,openright,blackthesis]{phdthesis}`. It changes the `papercolor`, `textcolor`, and `link-,cite-,url` colors. It always changes the figures path by appending a `blackfigs` to it. This means that a whole new set of figures are used to provide the user with both options.

The cover and back pages are likewise changed to match the `blackthesis` option and it allows for matching cover images.

The colors can be configured in the `PHDTHESIS.CLS`

1.2 PREAMBLE

The directory preamble consists of four files. I inherited Listing.tex and misc.tex and I added drafts.tex and misc.tex myself.

misc.tex

Define handy shorthand commands for math and latex variables. This is handy for quickly changing the appearance/notation in a quick and consistent way.

draft.tex

draft.tex is a bunch of draft package like todonotes that allows for missingfigures, margin notes and inline notes. These can then be viewed in the todolist which is sort of like a table of contents for todos.

Having these in a separate file that can be commented out at the time of final compilation should ensure that no todos or blindtext can be left in the final document without raising errors.

Figures

If you know that you will have a figure of something, it can often be handy to include a placeholder. This can be done with the missing figure command `\missingfigure` or the `includegraphics` command using a placeholder, `\includegraphics[width=\linewidth]{example-image-golden}`.

1.2.1 statics.tex

This file contains the information on the front and back pages. The information here is used in the file FRONTMATTER/FRONTPAGE.TEX (described in section 1.3.1) and in BACKMATTER/BACKPAGE.TEX (described in section 1.4.1)

This is implemented an official DTU template that I don't recall where I found.

1.3 FRONTMATTER

Here one could add: Dedication, abstracts, acknowledgements, notation, preface, list of publications, acronyms, etc.

1.3.1 frontpage.tex

Adjust as you please

1.3.2 contents.tex

Contains toc, list of figs, notation,

1.4 BACKMATTER

This could contain: Appendices, bibliographies, colophon (extra last page), publications.

Note that the blackoption might not work too well with inputting pdfs. You would probably have to manually set the background color of the page for these pages. You can use the color names defined in the .cls to keep it dynamic.

1.4.1 backpage.tex

Text can be added here to describe the project.

1.4.2 A bit about references

The reference can be set up in the class file.

The usual ref command gives: ??, 3.1 and 1.4.2. But in the misc file, we define special reference commands for each depth that includes the prefix: chapter ??, section 3.1 and section 1.4.2.

Likewise, we define other robust reference. The command can be seen below.

Chapter 1. Documentation of class

```
\newcommand{\partref}[1]{Part \ref{#1}}
\newcommand{\figref}[1]{Fig.\ \ref{#1}}
\newcommand{\tabref}[1]{Table \ref{#1}}
\newcommand{\secref}[1]{\textsection\,\ref{#1}}
\newcommand{\chapref}[1]{Chapter \ref{#1}}
\newcommand{\appendixref}[1]{Appendix \ref{#1}}
\newcommand{\listingref}[1]{Code Listing \ref{#1}}
```

There are backreferences for the citations.

MY WRITING PROCESS

2

I leave my brief thoughts on my writing process.

2.1 THE BRAINDUMP

Hopefully, throughout my time as a student I have been doing some work. Writing papers, notes for yourself, code, making plots. Collect these. Group them.

2.2 THE OUTLINE

Based on the groups. Try to create a coherent story that include all of your groups. This might not be the chronological you did the work. Use bullets, and figures. Now what should go in each chapter, section, subsection and paragraph.

Discuss this with my supervisor. It is a lot easier to adjust things now rather than once I have finished all of the writing.

2.3 THE VOMIT DRAFT

Get the thoughts out of your head and onto paper. It is so much easier to edit than to start from a blank page.

2.4 TREES MAPS AND THEOREMS

Meta: This summarises this book in very short.

This class is built on this book by Jean-Luc Dumont¹. It defines three laws

1. It treats different aspects of scientific communication but here we focus on written documents

FIRST LAW

Adapt to your audience

My audience is first and foremost my committee. Next, I would like if master and Ph.D students might benefit from reading certain parts of the thesis. In either case I find it important to write in a style so I myself would have been able to understand at the beginning of my PhD studies. This is my starting point. With that being said, I have been more thorough or gone to a lower level in the part on geometry than for my own work. My intention is that the geometry chapter might be used for learning the elemental differential geometry needed for manifold learning using Einstein notation.

SECOND LAW

Maximize the signal-to-noise ratio

Tried doing this by not going into too much detail on methods that I expect people to know. I also have a matetext at the beginning of each part, chapter and section that explain the aim of the writing. I have tried following Tara Brabazon's ideas about topic sentences and the paragraph which lead me to develop a certain structure for each paragraph, see more on this in section 2.5

THIRD LAW

Use effective redundancy

Read

2.5 THE PARAGRAPH

Topic sentence, bridge sentence, template for a paragraph, one idea per paragraph. [Link to Tara](#).

Tips: 1. Construct a strong, clear topic sentence. The topic sentence should identify the main point of your paragraph. As a general rule, topic sentences should be clear enough that a reader can get the gist of your paper just by reading the topic sentences of each paragraph. Try to keep topic sentences simple (e.g., avoid breaking them up with commas). Once you've written your paragraph, it's helpful to go back and check that a) you have a topic sentence, and b) it clearly captures the focal point of the paragraph.

2. Each paragraph should make one main point. In general, try to keep paragraphs between 3-5 sentences. If your paragraph is getting too

long, it is probably making more than one main point, and it may be time to break it into two (and make a new topic sentence).

Main idea: ... Old info: The manifold assumption is age old and use for justifying a lot of research. I New info: Another way to put this into use is to exploit/compute the geometry explicitly by learning the metric and computing geodesics

Gaussian processes are considered too smooth and therefor does not capture the geometry.

3. Establish internal flow by placing old information first and new information last.

2.6 SANITY CHECKS

This is a list of my sanity check, I would have liked to go through before submitting. I leave it here as it might be useful to someone else.

- All figures, tables, definitions, etc are referred to
- All citations work, no ??
- All references work, no ??
- All figures have a label
- All figures should have a short name
- All figures should be placed properly, especially marginfigures. antimjustification is needed.
- Clearly state what is my work/contributions/novelty
- Check that all acronyms are added to the Acronyms
- Check that everything is added the index
- Check that no todos or example images or missing figures remain
- Check for consistent notation
- Check that notation page is up to date
- Check that color schemes work in both black and white thesis options if you are using the black options
- Remove draft packages

Chapter 2. My writing process

- Remove half sentences starting at a new page (in Danish: hore-unger)
- Decide on citation style
- Spell check
- OMG gammar
- and the dreaded commas

RANDOM EXAMPLES

3

The manifold assumption is central to my work: We assume that high-dimensional data lives on a lower dimensional manifold embedded in the high-dimensional space. Quite simply, this thesis explores how to learn this manifold. Understand the manifold assumption intuitively thoroughly. We won't focus too much on traditional methods but rather tie this in to probabilistic methods.

This is all just a collection of random examples

3.1 THIS HERE IS A MUCH LONGER TITLE FOR A SECTION BUT THE SHORT TITLE SHOULD BE SHOWN IN TOCS.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.



Figure 3.1: This figure should be aligned with the text top

3.2 OTHER EXAMPLES

Table 3.1: Minimum Requirements¹ for Automatic Readmission into the Commerce Faculty

| | BCom | | B.Bus.Sci | |
|------------|------------------------------------|-----------------------------|-------------------|-----------------------------|
| | Number of courses required to pass | Cumulative Total of Courses | Number of courses | Cumulative Total of Courses |
| First year | 4 | 8 | 4 | 18 |

¹. This is a footnote in the margin - different from a marginnote

Figure 3.2: Illustration that any curve can be discretised and the lengths of the increments can be computed using Pythagoras. See if this figure can be merged with Tosi et al, 2014 figure 2. It might have to be full width.

Golden ratio

(Original size: 32.361×200 bp)



Full pagewidth figure to illustrate the idea of the VR complex: figure of points, points with balls with small d , points with balls with large d

Figure 3.3: illustrate the idea of the VR complex: a) figure of points, 2) points with balls with small d , 3) points with balls with large d

LAMBERT-BEER'S LAW

$$I(x) = I_0 e^{-\mu x}.$$

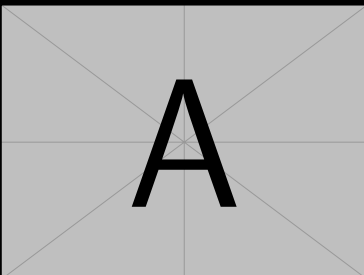


Figure 3.4: Attenuation for all the different processes for Carbon in the energy range 5 – 150 keV. T for total, P for photoelectric, C for Compton Scattering, and R for Rayleigh Scattering (Coherent Scattering). From [XRayOpticsCalculator]

3.3 REMARKS AND THEOREMS

Unnumbered theorem-like environments are also possible.

Theorem 3.1 *Let f be a function whose derivative exists in every point, then f is a continuous function.*

Theorem 3.2 (Pythagorean theorem) *This is a theorem about right triangles and can be summarised in the next equation*

$$x^2 + y^2 = z^2$$

And a consequence of theorem 3.2 is the statement in the next corollary.

Definition 3.1 *There's no right rectangle whose sides measure 3cm, 4cm, and 6cm.. I have numbered definitions with respect to chapters*

You can reference theorems such as 3.2 when a label is assigned.

Remark 1 *Given two line segments whose lengths are a and b respectively, there is a real number r such that $b = ra$. Remarks are not numbered*

Definition of something random

HEJ

Just learnt how gather in the amsmath package can be used to align equation AND have text in between. OMG!

$$\mathbf{F}_{\text{net}} = m\mathbf{a}$$

This is the Newton's second law of motion when the mass does not change with time. But for time-varying mass, we have to use

$$\mathbf{F}_{\text{net}} = \frac{d(m\mathbf{v})}{dt}$$

3.3.1 Fancy Quotes

“ *That, if a straight line falling on two straight lines makes the interior angles on the same side less than two right angles, the two straight lines, if produced indefinitely, meet on that side on which the angles are less than two right angles.* ”

—Euclid

“ *Where there is matter, there is geometry.* ”

—Johannes Kepler

3.3.2 Fun little thing

Aaaaaaargh he shouted but not even the next one in line noticed that something terrible had happened to him. Could maybe use this in introduction/preface or something.

3.3.3 Options in the blindtext package

1. First item in a list
2. Second item in a list
3. Third item in a list
4. Fourth item in a list
5. Fifth item in a list

First item in a list

Second item in a list

Third item in a list

Fourth item in a list

Fifth item in a list

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{i=n} x_i = \frac{x_1 + x_2 + \dots + x_n}{n}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\int_0^{\infty} e^{-\alpha x^2} dx = \frac{1}{2} \sqrt{\int_{-\infty}^{\infty} e^{-\alpha x^2} dx} \int_{-\infty}^{\infty} e^{-\alpha y^2} dy = \frac{1}{2} \sqrt{\frac{\pi}{\alpha}}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\sum_{k=0}^{\infty} a_0 q^k = \lim_{n \rightarrow \infty} \sum_{k=0}^n a_0 q^k = \lim_{n \rightarrow \infty} a_0 \frac{1 - q^{n+1}}{1 - q} = \frac{a_0}{1 - q}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{-p \pm \sqrt{p^2 - 4q}}{2}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} + \frac{\partial^2 \Phi}{\partial z^2} = \frac{1}{c^2} \frac{\partial^2 \Phi}{\partial t^2}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression

Chapter 3. Random Examples

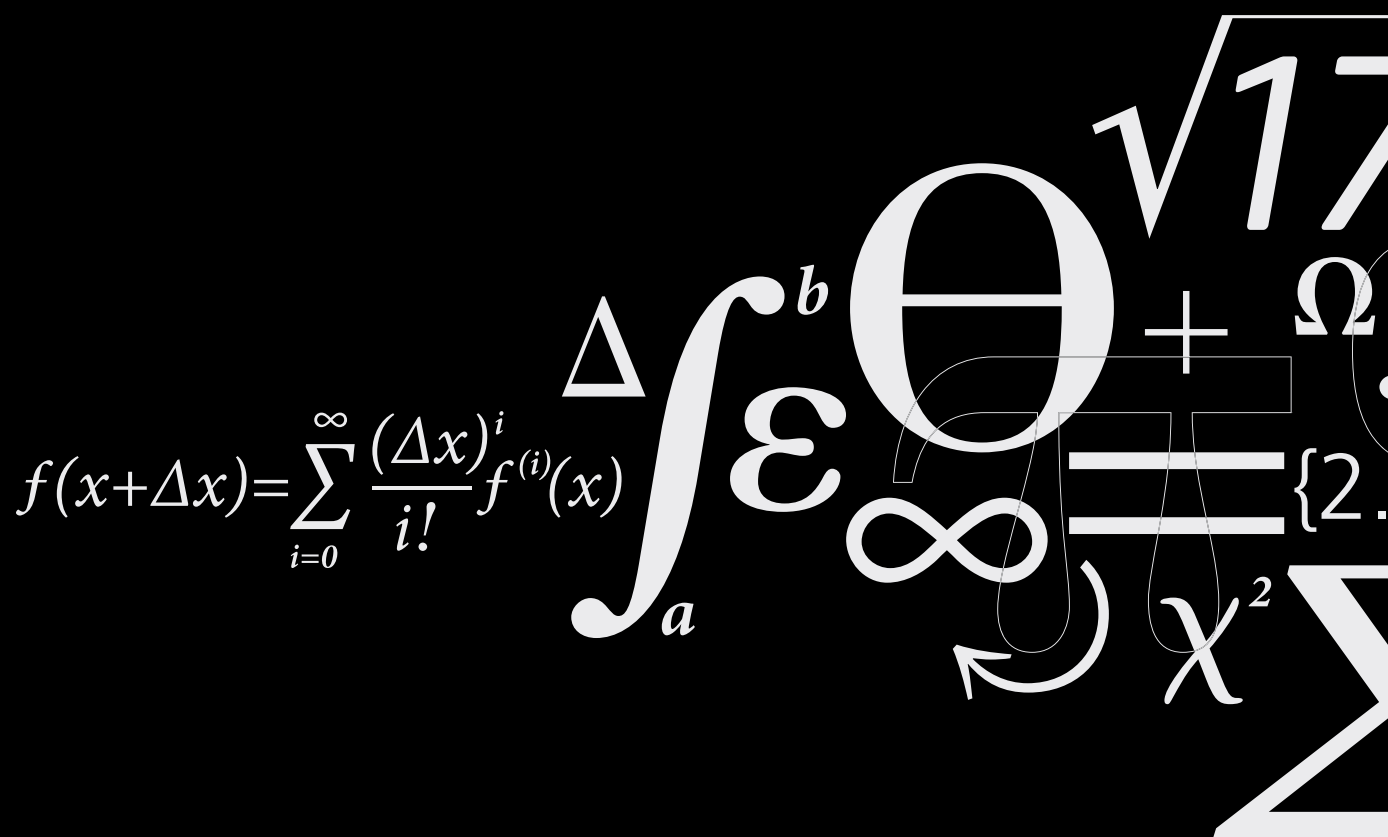
of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

EXPERIMENTS

4

Github branches, colab notebooks?

APPENDIX



TECHNICAL UNIVERSITY OF DENMARK
COGNITIVE SYSTEMS

RICHARD PETERSENS PLADS, BUILDING 321
2800 KGS. LYNGBY
PHONE +45 4525 1700

WWW.COMPUTE.DTU.DK