

# Can we date backdated paintings using Data Analysis?

Author: Molly McCarthy

Final Year Project – BA Digital Humanities and Information Technology

Supervisor Dr. Andrea Visentin

Discipline: Digital humanities

Institution: University College Cork

04/2025

**Abstract**

This project contends to investigate whether data analysis and visualisations to discover backdated paintings within an artist's oeuvre. It attempts at developing a method of authentication using artefact metadata using the inventory dataset of a museum. The entire dataset is composed of a single artist's work. It enhances the efficacy of these methods by integrating biographical data surrounding the artist and the context of the life they led.

### *Declaration of Originality*

*In signing this declaration, you are confirming, in writing, that the submitted work is entirely your own original work, except where clearly attributed otherwise, and that it has not been submitted partly or wholly for any other education award.*

*I hereby declare that:*

- *This is all my own work, unless clearly indicated otherwise, with full and proper accreditation;*
- *With respect to my own work: none of it has been submitted at any education institution contributing in any way to an educational award;*
- *With respect to another's work: all text, diagrams, code or ideas, whether verbatim, paraphrased or otherwise modified or adapted, have been duly attributed to the source in a scholarly manner, whether from books, papers, lecture notes or any other student's work, whether published or unpublished, electronically or in print.*

Signed: Molly McCarthy

Date: 21/04/2025

### **Acknowledgements**

Thank you to the FelixArt Museum, the Royal Institute of Cultural Heritage of Belgium and to Dr Brynn for allowing me into the fold.

I would like to thank Dr Andrea Visentin, for always showing me the way forward, and accommodating me when I really needed it

I would like to thank Dr. Shawn Day and the entirety of the UCC Digital Humanities department for supporting me on this journey.

# Introduction

In this project I investigated if we can discover potentially backdated paintings within an artist's oeuvre using data analysis techniques. The artist I investigated was Felix De Boeck, he was a Flemish Abstract painter in the 20th century living in Belgium. I am undertaking this project in collaboration with the FelixArt Museum and the Royal Institute of Cultural Heritage of Belgium (RICH). I have been provided with the inventory spreadsheet of the FelixArt Museum, which houses a large quantity of De Boeck's work. Investigating his work was a task assigned by the Museum. A secondary task I undertook is the development of a methodology using the physical descriptors of a painting using data analysis. I completed this task using Python's Pandas Library and Google Colab for data processing and visualisation. My commitment to the Museum is to return my findings to them.

Backdating is the practice of putting a date on a painting that is much earlier than it was created. This is usually done for monetary gain. Felix De Boeck (1898- 1995) was a farmer for most of his life. He painted consistently from 1915 to 1989, and the FelixArt Museum which houses a large amount of his works is constructed on his land in Drogenbos, Belgium. Using data analytic methods I search for features across the dataset that are anachronistic to its time period in order to discover potential leads to authenticating the collection.

## Literature Review/Environmental Scan

In this literature review, I briefly discuss the European art market of the 20<sup>th</sup> century and De Boeck's position in it. David and Oosterlinck analyse the market during and after the second World War, and why De Boeck could have backdated his paintings. Guerra presents a history of archaeometry. In the second part, we investigate approaches to detecting forgeries using machine learning. Polatkan et al. provide a literature review of such approaches.

Finally, I examine two pieces from Windhager's Doctorate dissertation: the first studies of what types of visualisation techniques are used for Cultural Heritage collections, and the second about how Biographical data and Collection data can be displayed on the same visualisations.

## War. Inflation, Monetary Reforms and the Art Market

- Geraldine David, Kim Oosterlinck

Belgium was occupied by Germany during the second World War in 1940 and was liberated in October of 1944. The art market in occupied countries during this period noticed a massive increase, and this paper sets out to find out how this may have been impacted after the liberation, and what exactly were the motivations of the buyers in this period. This study first assembles a history of the economy during this period and the following years to examine what was happening with regards to the art market given the extreme circumstances and immense controls placed upon the economy by the invading state. Then they examine the record books of the Galerie Georges Giroux during the period of December 1944 to May 1951, no data is available from during the occupation as the gallery was not in operation. They find that there was a large crash in the market immediately following the liberation, which they attributed to the controlling of the money supplies and public distaste towards anyone seen to be profiteering from war following the occupation. The market slowly increases again towards the end of the decade. They determine that the purchase of art was not a method of investment, as after the liberation of Belgium investors did not rebalance their portfolios toward stocks. The increase in the purchases of art while it cannot be pinpointed during the wartime period is most likely attributed to a form of money laundering by shopkeepers and black market dealers who were making massive profits during this time, and the art market was one of the few ways of divesting their money that was not overseen.

This information provides another angle for my research, as we are attempting to determine when Felix De Boeck may have been backdating some of his works. We know there was massive demand on the market during the occupation from this paper, so the chance that he may have been painting more to sell more is possible. This alone may not inform his behaviours during this period, as the public's views on the shopkeepers and black-market dealers capitalizing on the increase in demand during war was less than favourable. I will speak more on how this informed my research in the Methods section.

## Archaeometry and Museum: Fifty Years of Curiosity and Wonder\*

- M.F Guerra

Guerra presents a history of Archaeometry and how it has evolved alongside museums. Archaeometry plays a significant role both to learn more about the history of artefacts in museums and to discover if they are legitimate. Many old and ancient items were first kept in personal collections, which over time were donated to or became museums. Due to these items being overseen by collectors rather than archivists or scholars, the amount of historical information regarding the object is lacking. Furthermore, as the desire to collect these items was simply to have them and to show them off, many opportunists began to produce fakes. This paper shows how museums and scholars learn about artefacts across different mediums and materials, with a particular focus on detecting fakes. A vital use of archaeometry is also in the determination of materials and techniques used to create artefacts such that they can be preserved and restored effectively. One such case was in the restoration of some of Gustave Moreau's black carbon drawings, through the use of elemental and structural analysis alongside testing them under numerous types of lights they were able to determine the differences between graphite, charcoal and black chalk, enabling them to be correctly restored. Guerra also points to how these scientific techniques are used to

authenticate artefacts and discern them from fakes, but also how the forgers were influenced by these techniques. Forgers looking to pass off fake terracotta figurines in the late twentieth century were breaking down legitimate terracotta from the time periods they were attempting to fake and incorporating them into their productions in the same places samples were typically drawn from.

For my investigation attempting to discover backdated paintings in De Boeck's oeuvre it's important to understand what methods are currently being used in order for museums to authenticate items. While many of these methods are ways of finding genuine evidence to prove and authenticate these items chemical analysis, something that cannot truly be ascertained with analysis of physical metadata, the lessons learned can be applicable. Of particular note is that the production of forgeries is often opportunistic, and aware it's under scrutiny. Guerra mentions toward the end of this paper that not only do the materials matter in the quality of a fake, but also the ability of the faker; in the case that De Boeck's oeuvre contains backdated works it's most likely he did them himself, many Archaeometry techniques will be needed to determine their authenticity.

## Detection of Forgery in Paintings using Supervised Learning

- Güngör Polatkan, Sina Jafarpour, Andrei Brasoveanu, Shannon Hughes, Ingrid Daubechies

Whilst my analysis is centered on the anachronisms of how an artist treated his works overtime, this paper directly addresses the brushstrokes on paintings and the colour representation. They use this as a method of authentication using Machine Learning. The article is centered around discoveries made at the First and Second International Workshops on Image Processing and Artist Identification (IP4AI). They used a dataset consisting of images of original Van Gogh paintings and images of copies. For the Second workshop they expanded the dataset to new originals and fakes. Using their previous techniques to separate originals and copies they expected them to work, but they ended up grouping every new item

in the dataset together, both originals and copies. Further investigation revealed that there was less blur on the new images as the scans were completed using a new acquired higher quality scanner. Blur was disrupting the hesitancy measurement they were conducting on the paintings, essentially testing how many micro-movements were made with the brush as less micro-movements meant a more confident stroke. In the first workshop hesitancy had shown to be higher for fakes on average than originals. However, the hesitancy score fell dramatically across the board for all of the paintings with sharper scans, hence why the model was failing. In order to solve this issue, they decided to develop a new method of testing, introduce new features to evaluate and develop a new dataset to build their model on. They began tracking the coordinate representation of colors using the original RGB. This will help them determine differences between multiple paintings. They also refined their stroke testing algorithm. In order to normalize their dataset, they developed new paintings and new copies very quickly together, all painted by the same artist. This was done to make the dataset more difficult to discern in order to improve the model. It also reduced the number of differences in the materials used between the original and copies immensely, decreasing the discernability again. They ran three tests on this dataset, and it showed promising results for the first two tests but went noticeably down on the third test where they switch from oil paints to acrylic paints.

For my research project, this paper shows the significance of certain physical properties of paintings that are not within my dataset which may prove to be useful. Performing their analysis in the same ways brushstrokes and colours could not be performed to the same extent for a visualisation style approach, as opposed to their use of machine learning, but perhaps they could be developed for use in the future with this method. The part of their discoveries which interest me the most is use of a “ground-truth” dataset, the one they developed themselves for training the method. If I could understand exactly what a backdated



painting looks like within my analysis would help me to find other anachronistic paintings within the dataset.

## Visualization of Cultural Heritage Collection Data: State of the Art and Future Challenges

- Florian Windhager, Paolo Federic, Günther Schreder, Katrin Glinka, Maian Dörk, Silvia Miksch, Member, IEEE, and Eva Mayr

This paper is an exploration of the use of Visualisations on Cultural heritage data of GLAM institutions (galleries, libraries, archives and museums). It examines how and why a user might use them, what typically gets visualised, and how things are chosen to be visualised. They used a dataset of seventy different visualisation sets across research papers and standalone online implementations. Most of these visualisations are created to aid the user in their information journey, be it by better representing the data and making it easier to understand, by authoring a path of thinking for the user to follow, or just to make it nicer to look at than a standard spreadsheet of data. They find that collections are graphed differently depending on the user groups, which lines up with the types of information journeys the groups are on; casual users generally browsing through data, whereas expert users will be searching. As such there is a higher proportion of visualisations designed for expert users represented in research papers than on public website implementations.

This paper is quite dense and to summarise it fully would be beyond the scope of this literature review, so I will continue by only summarising what I deem to be relevant for this project. A total of 81% of visualisations chose to visualise the temporal access as a 2nd or third dimension. As the goal of this project is to ascertain anachronisms as they appear in relation to time, it is reassuring that my audience will find this easy to understand as they are commonplace among Cultural Heritage Visualisations. It is important to note that this paper does not look at visualisations specifically used for Data Analysis where the visualisation

aids the investigation, rather than the visualisations are used to better communicate their findings to the user. This paper also shows that there is a precedence for the representation of cultural heritage data as visualisations.

## A Synoptic Visualization Framework for Artwork Collection Data and Artist Biographies

- Florian Windhager

The above paper is part of this PhD Thesis. As such the focus of this segment of the literature review will be on its fifth section, where it synthesizes together the sum of the above paper and three others to investigate further into the mapping of biographical data and collection data together.

First, they clarify that most biographical data is tracked as an event collection, typically in a linear fashion that maps point of information with regards to events that occur within a person's life. Whereas data regarding a collection is more systematic i.e. object oriented. This means that typically graphical representations combining these two forms of data will lend themselves towards being visualised along an event collection, which is an arbitrary timeline based upon the biographical data at the disposal of the study the visualisation is being made for. This is not necessarily the case for some macro-level, where you would visually represent a field of systems that would visualise the greater context surrounding the collection. This might be one specific time period that looks at the factors at play that may have affected the creation of that collection i.e. economy, pop-culture, politics etc. With regards to mapping biographical and collection data along time, timelines are quite efficient. Either multiple one-dimensional timelines on each collection and biographical data that then contextualise each other, or graph-time-cubes. The latter of these tends to map temporality vertically and can place events within another axes of information. Most interesting of these is for geo-temporal mapping, where the depth axes of the cube are a set of

coordinates over a map, visualising the time and place of events i.e. location of a photo and the year it was taken.

This paper offers a guide as to how one would best marry biographical and collection data into visualisations together. Based on the lack of rich biographical data I have relating to Felix De Boeck due to a lack of a translator, many of these visualisations are beyond the scope of this project, yet it's principles on how to link biographical and collection data will prove invaluable.

## Tools

The selection process for the tools employed in this project was informed by two main criteria: ability to complete the task assigned, and accessibility to develop an accessible methodology for future use. The data was provided by the RICH and the FelixArt Museum and necessitated its use as part of the project proposal, therefore the tools that I selected for this project exclude the choice of dataset. The tools selected must be able to efficiently and effectively be able to process and manipulate the data, enable the creation of detail rich visualisation from which analysis may be drawn, and an effective means to create a digital artefact. The secondary goal of creating a methodology framework that can be used by other researchers/institutions brings the accessibility and ease of use of the tools into consideration.

The dataset was provided in the form of a Microsoft Excel spreadsheet which can be exported into different formats, thus not limiting the range of tools that can be implemented. The dataset is the inventory sheet for the FelixArt Museum's collection, so its primary use is for data collection and tracking rather than analysis. To use the data, it must be processed and cleaned before visualisations can be created. To accomplish this task, many tools were taken into consideration: OpenRefine, Python's Pandas Library, Tableau, and PowerBi.

OpenRefine is a powerful open-source tool for operating on data, particularly cleaning and preparing data for other uses. It has a list of powerful features such as faceting, reconciling and clustering. All of these can be used to clean, group and truncate messy data into a usable format, which can then be exported into most any format that you might require. Facets allow you to categorise and map your data into whatever layout you need quite easily. This would be especially useful when it comes to the large amount of categorical data within my dataset. Clustering and Reconciliation enables you to clean away messy data, clustering via OpenRefines heuristics system, essentially being able to consider all the values in a column and offer the best option to consolidate them into the most concise number of values. Reconciliation is quite similar, but differently to clustering it bases its options to reconcile the data on external sources i.e., wikidata commons. This can be useful for consolidating categorical data like names or professions if these are misspelled a resource like wikidata probably offers something quite close. This is useful if you are integrating the data you are cleaning to an external source, which is not something this project is considering. It is also important to note that this tool can only perform data cleaning and preparation, none of the other tasks required for this project. If completing the task assigned by the museum was the only criteria this would not weigh heavily in the selection process, but my secondary goal of creating a methodology would value a more integrated system, as reproducing the process would be a more compact endeavour.

Pythons Panda's Library is an extensive suite of tools built on top of the Python language. It is built upon a number of other libraries like Matplotlib and NumPy and integrates well with base Python functionality. The core of Pandas is built around the dataframe object. It is a multi-dimensional data structure that handles data in series and rows, similar to a spreadsheet. Data can be read into this object using many common data formats like csv or tsv files. Python data manipulation methods for standard python data structures are

overloaded to work with dataframes as well as new methods included with Pandas that allow for easy and effective data cleaning and processing. The ability to copy dataframes and manage multiple at a time allows for easy processing for many applications. Pandas is built with Matplotlib in mind and thus can also be used to create visualisations. All of this means that this suite of tools would be easy for anyone with programming experience to use, and due to the reusable nature of code would make my process quite easy to replicate.

Tableau is a data visualisation and business intelligence software. It is comprised of a suite of tools necessary to process data, create visualisations and generate reports. The Tableau Prep Builder tool is the system used for data cleaning and preparation. It enables you to do most any cleaning and structuring method you would need. The central part of Tableau Prep Builder is the Flow system. It is a graphical user interface that tracks all the processes you have conducted on your data, while also allowing you to reverse them if needed. New datasets and processes can be dropped into the flow to add to it. You are also enabled to look at the data you are processing using a Profile window, as this shows you the overall structure of the data. This pane is where smaller fixes and operations happen such as renaming columns or fixing errors in data entry. Finally, there is a rows tab which shows you a spreadsheet view of the data you are operating on, where you can perform cell level operations if you want. Tableau can also generate rich visualisations, so it represents an inclusive package which is preferred for this project. This can cause access problems, primarily in terms of cost. The license is currently sitting at €900, which does not account for time to train an operator or investing in a system capable of running it if required.

PowerBi is much the same as Tableau. They are both popular in industry, and highly capable at processing data, creating visualisations, and generating reports. PowerBi, as Microsoft creates it is quite similar in layout and visual design to the Microsoft Office suite, so most users familiar with those programs would find a familiar experience. The steps

window in PowerBi tracks the different processes conducted on the data in a list-like structure, similar to the Flow tab in Tableau. Refining the dataset is similar to operating excel, applying filters and formulas to modify the cells accordingly. To create visualisations, you can select from a long list of available styles of charts and graphs. Once your data is loaded into this system, you can tweak and change how the chart is laid out. Similarly to Tableau, PowerBi is a paid software, but it is a lot cheaper. A single user license is €14 for the baseline package if you wish to be able to share/export the visualisations you create. It is also possible that potential operators of this projects methodology would already have access to the software via Office 365 as some tiers of that product include a PowerBi License.

I have touched on some of what these programs are capable of when creating visualisations. To briefly recap, OpenRefine cannot create visualisations and so would require use of multiple pieces of software, Pandas is fully capable of creating visualisations with granular control and both PowerBi and Tableau offer a curated suite of visualisations to choose from and modify to achieve the final result you desire. Due to accessibility concerns I will not be proceeding with Tableau or PowerBi. This leaves Python's Pandas Library as my best option for creating visualisation. As stated above, I have a preference for developing a streamlined workflow for working with the data and creating visualisations, so ultimately, I decided to use Pandas for this project. While it would be possible to use the terminal to create these visualisations, the modern comforts of an IDE would be quite helpful. Google Colab is a free, cloud-based platform that allows users to write and execute Python code in a Jupyter Notebook environment. It provides access to powerful computing resources, making it ideal for data analysis. It also includes its own Google Colab library for Python, which enables the transformation of standard dataframe outputs into interactive tables, which can be filtered and searched. An added benefit of this choice is that the notebook is hosted on a Google Drive, a

Cloud storage system which I can use to create a publicly hosted site that can house the digital artefact for this project.

I will briefly describe the process of transforming the raw data from the dataset into visualisations in order to sketch out the flow. The data is exported from Excel into an easily readable format for Python i.e. CSV. Hosting this in Google Drive Project folder for my Colab notebook and using a Pandas integrated methods this is read into a Data Frame object. The data is then processed within the Data Frame object using Pandas methods. Once this is complete, the data is selected and grouped as desired for the visualisations to be created. Once the data is in this workable state, you generate a visualisation in the style of your choice using Pandas methods, and then you can modify attributes to gain granular control over the outputs.

## Reflections on tool selection

Python and Pandas worked well, and the Google Colab plugin enabled powerful data investigation via its interactive datasheet option. Pandas did offer a large issue regarding its multiple methods of copying or modifying specific cells. While there are multiple methods of doing this, none of them work every time due to the ambiguity of these methods. This means that sometimes they would work, and sometimes they would not save your changes for further use. Even the recommended method on the community made documentation is not a surefire success in every scenario. This caused major delays and led to a lot of confusion when working on the project. This is something that could have been solved quicker with better error messages from Pandas and Python. This issue was also compounded by Google Colab's lack of a traditional debugging tool. Not being able to see the journey of a piece of data through my code through a debugger was incredibly frustrating and led to a major increase in the time spent on processing the data. Colabs most earnest solution to this is the integrated Gemini AI copilot's opinion on what the error message means and how to fix it;

the handful of times I used this option out of sheer frustration it did not provide working code. The final gripe I had with Google Colab is that it arbitrarily decides how much of your text output to display, limiting the amount of your dataframes, lists and series that are printed to the output. In the case of dataframes, Colab's integrated Python library can overcome this issue, transforming the output into an interactive table, but there was no such luck with lists or series. This disrupted workflows and made viewing the data I was working with difficult at times, particularly as I would be creating a series object just before creating a visualisation with the data I was looking to work with.

## Method

In this section I briefly discuss a scenario which explains how the physical description of the paintings may reveal anachronistic features of paintings in De Boeck's works. Then I describe the dataset and point out which features are of particular note and provide context as to why. Then I will highlight certain parts of biographical information I will introduce to the development of my experiments. I will then briefly explain the way in which this method could be developed to be used in further research, and particularly how it may be of benefit. I will then quickly reflect on what worked with this approach and what did not.

The core principles of this analysis are that we can link how De Boeck interacted with his paintings to his behaviours at a given point in time. For example, if he signed his name in the same place on his works for 20 years, but then he changed the position after this. What this would reveal in the case of a backdated painting would be one painting in a sea of other works from that time that is signed not only differently, but the same way he signed it at only another time in his life. Of course, this would be incredibly simple and removes a lot of real factors from the scenario, but it is representative of an anachronistic feature in a painting. This one feature alone would not be considered robust evidence in the case of a backdated



painting, even in such a black and white scenario as sketched. However, if we examine enough factors and features of his work, we might be able to build a strong enough case for certain paintings to garner further investigation by the museum with more traditional methods of authentication. Data Analysis paired with rich visualisations can enable us to investigate the data and compare distinctive features in a usable and communicable way.

## The Dataset

The dataset is the inventory spreadsheet kept by the FelixArt Museum in Drogenbos Belgium, built on the same land De Boeck lived and worked on. It houses a large collection of his works. It contains contextual metadata for the works, such as their title, creator's name and production year. The data set describes physical properties of the paintings, such as the size, materials used, how and where he inscribed his signature etc. It also includes a large amount of information regarding the maintenance and continued housing of the piece as a museums artefact, such as when it was obtained by the museum, processed into this spreadsheet and how much it costs to insure for example. There are 274 columns and 873 rows in this dataset. Each row represents an item kept in the museum. The data set is not entirely populated; some columns are entirely empty. This is important to note as it highlights that this dataset was not designed to be analysed, but to be used as a database of information to be referenced.

Interesting features about this database relates to how Felix De Boeck painted. A number of his paintings were painted on both sides of the canvas or were composited from multiple works. This leads to some paintings having vastly different dates for when their production began and when they ended that would stretch much more than just a year or so. It also means some of the works have multiple signatures and could not be signed on the back. As it is an operating museum, they do have a few artworks in their possession that do not bear De Boeck as their creator. We can assume that this is for a temporary exhibition, or it

has significance to the experience the museum is creating, regardless, it must be removed from my investigation as we are only concerned with De Boeck's works. As this Museum is in part of the Flemish region of Belgium, the dataset is predominantly in Dutch.

The features that are intriguing for our analysis are generally physical characteristics of the paintings and the year the painting was completed. The physical characteristics of most interest are:

- dimensions of the works in centimetres (H cm x W cm)
- the description of materials used i.e., oil on hardboard (olieverf op hardboard)
- the plaintext representation of the inscription's contents i.e. FELIX 1927
- method of inscription i.e., incised into paint, in pencil etc.
- location of the inscription, typically described in zones i.e. top right, middle-left, back.

And why are they interesting? It was briefly mentioned in the introduction to this section, but I deem some of these characteristics to automatic decisions he may have made when sitting down to paint or finishing on that particular piece. Two things pointed out by Guerra (2008), one that one of the most important factors in a forgery is the ability of the faker, needless to say Felix was highly skilled at producing work that looked like his own and secondly that forgers are aware their work might be investigated; it's very possible that he used the exact same type of paint or the same type of brush, though not necessarily, yet he used many different signatures over the years for example. Something like a signature, or the size/type of canvas he had on hand could have been decisions that were influenced by other factors in his life, not the pure desire to backdate a painting. I must acknowledge that it is entirely possible that he did account for everything, and we might not be able to tell without

element analysis or a more concrete authentication method. Yet, that is what we are here to test.

Data Analysis is allowing us to step back and observe De Boeck's entire oeuvre from a distance and look for trends that appear across works. These can be trends within specific styles, types of signatures, sizes of paintings or trends across time. In a way we are attempting to determine his style, not his painting style but his style of working on a painting. It is a bit of romanticization, looking for the fingerprints he did not know he left across his paintings, but it might work. Particularly Guerra noted that forgers responded to changes in the space of Archaeometry, and it has advanced quite a bit since De Boeck may have been backdating paintings.

Like any crime, he would need a motive. This is difficult to pin down, particularly due to the lack of access to a translator over the course of this project. But there are some points of reference to draw from. Veltman's summary of his biography on Felix De Boeck that was translated into English, and the general economic state of Belgium and the art market over his life is where I draw most of my biographical influences on why and when De Boeck may have backdated paintings. During the early 20th century following the First World War, his paintings became quite famous, before he fell away into obscurity during the 1930s. He enjoyed a resurgence of interest in the late 1950s and early 1960s as he had been the subject of two pieces of work by Jan Walravens, one of which he hastily wrote on his deathbed. David and Oosterick (2012) do mention that the art market was a way for shopkeepers and black-market dealers to invest their money during the occupation, and Veltman points out that Felix already had a proclivity to selling paintings in private. I must acknowledge that there is more information which may suggest whether he would have interest in selling paintings around this time. His farm was successful enough that he did not have to make a living from his paintings, but also the economic situation in Belgium during the occupation

could drive someone to anything. His multiple portraits of Hitler, one in 1942 and once again in 1946, and association with fascists makes pinning down his interests during this time difficult. For the purposes of this investigation, I will be taking interest in paintings that are dated for early periods in his career, and that may bear some resemblance to his works during the 40's and onwards.

## Development of this method of investigation for further use

The main reason I feel this methodology can be expanded upon and used as an investigative tool for GLAM institutions to authenticate artefacts is because this dataset is quite simple in concept. As it is a catalogue of items maintained at the museum, the parts of this dataset that are interesting are not difficult to acquire, as opposed to other methods of authentication that may require scanning, x-radiography, or chemical analysis to name a few. Not to mention the massive costs both in time and money involved with these methods. If employed alongside some metadata standards, datasets of artists' work could be assembled even when the artworks are not in one singular location like the FelixArt Museum.

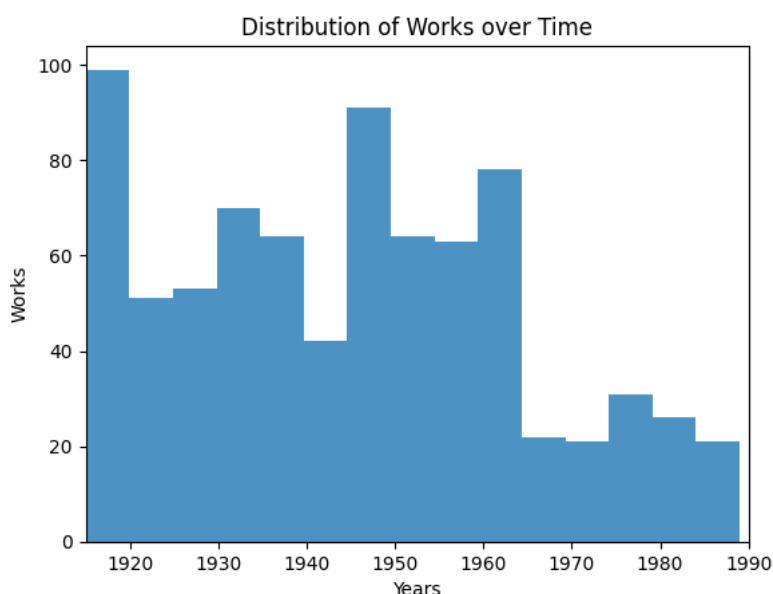
I will briefly describe my intentions regarding the method of Data Cleaning. This will only extend to things that are specific to this data set, the standard fare of data cleaning, such as reconciling issues in data entry, removal of empty columns etc. will not be covered in this overview but are involved in the process. The dataset includes some irrelevant data, this can include paintings within the collection that are by a different artist, or works that are in different mediums than which we are interested in for certain tests, such as a circular canvas when trying to measure the dimensions of the rectangular majority of his works for example. I quickly removed a number of columns from the dataset I was working with as there was a large number of columns that would not be useful in my analysis. Often when he signed his works, he would sign and date them. He had a double-digit number of common signatures

that appear across his oeuvre, and collating them together has required stripping away years and creating a new column.

## Analysis and Reflection

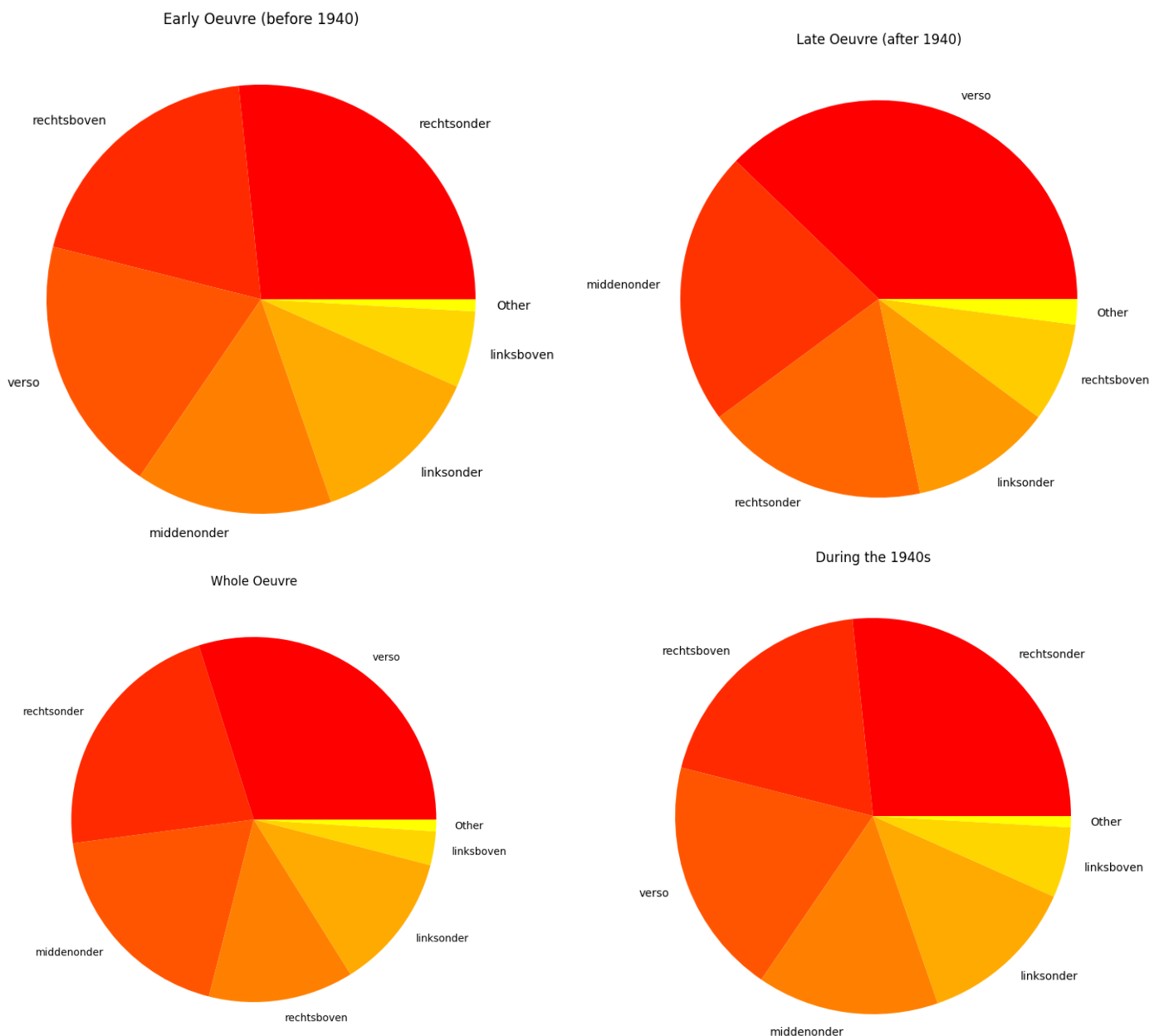
In this section I will walk through the development of my investigation as time went on, making reference to visualisations created throughout the project. I will then speak about what results I drew from this process. Then I will speak briefly about how I achieved regarding the creation of a new methodology for Authentication.

A consistent theme based upon the biographical information we can now, the most likely time period I can determine for when Felix De Boeck may have been backdating works would have been between 1940 - 1960, but with a specific interest in the 40s. I believe the period in the forties is likely due to the state of the art market and it's opportunities relating to the occupation of Belgium as well as the liberation immediately after. Following the turn of the mid-century, the work early 20th century European modernists and the rise of abstract expressionism in the states caused the value of the genres to increase (Robson, 1985), which may have incentivized De Boeck to cash in as the Belgian currency was volatile following the effects of the occupation on its economy (David and Oosterlinck, 2012).



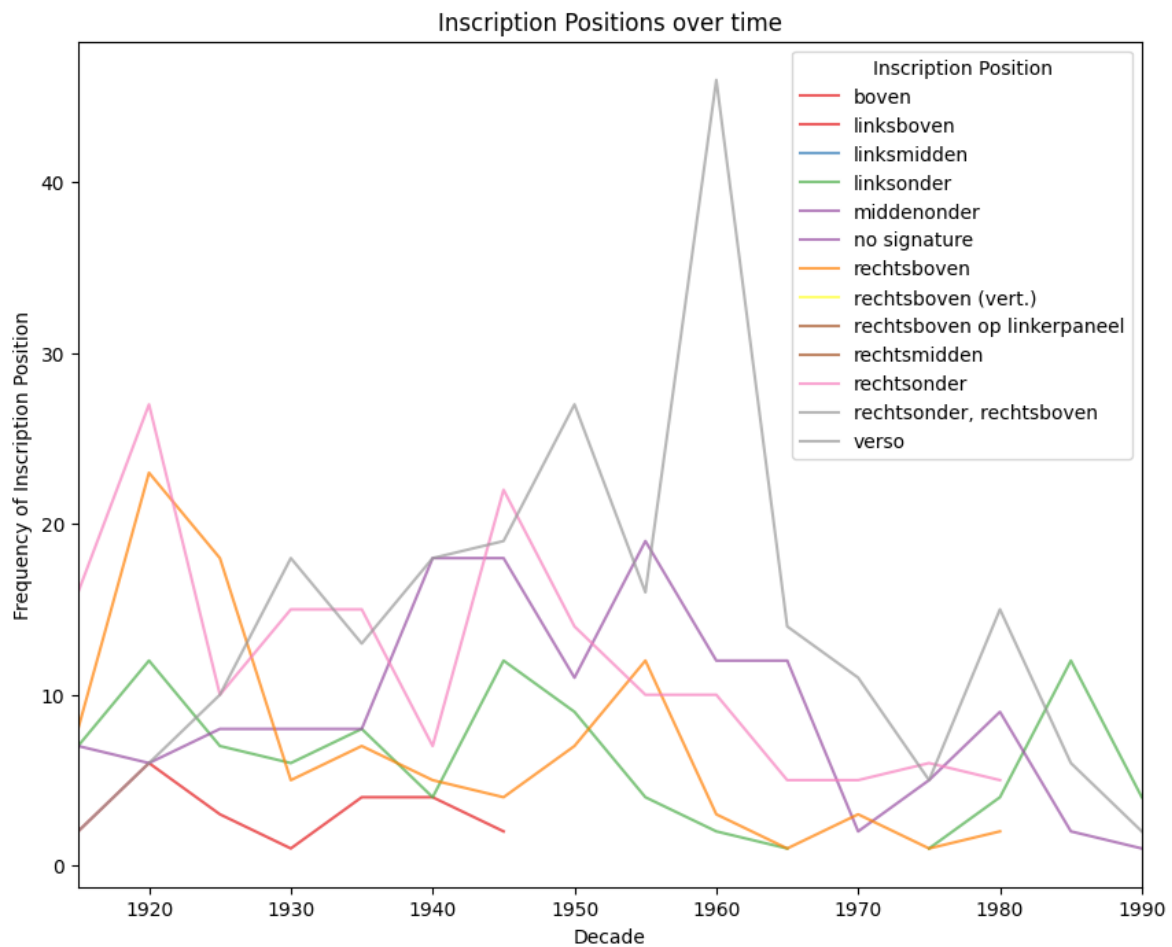
I began by looking at the complete collection of paintings, to get a feel for the lifespan of the man's working life. We can see in the histogram that he painted for much of his life. He painted the most when he was younger, during the tail end and just after the First

World War. It was during this period and shortly after where his works were held in the highest regard, and he enjoyed the most fame Veltman (2021). The next most noticeable point in this timeline is while Belgium was occupied during the second world war, and the sharp increase in the 5 years following the liberation. He painted frequently until the second half of the 1960's when he stopped producing works at the rate he had the rest of his life.



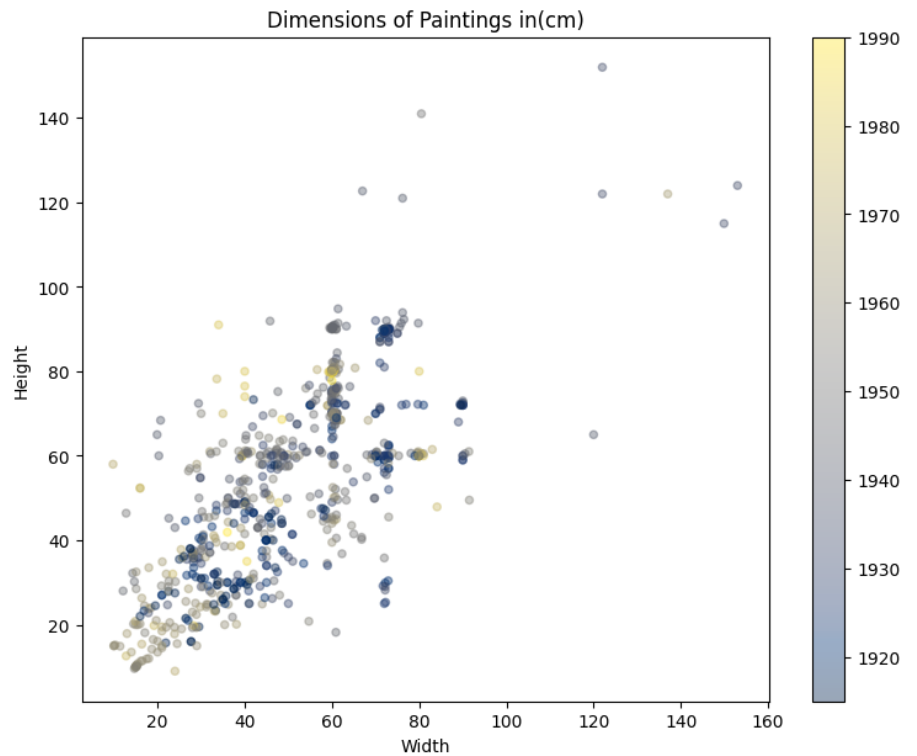
Dr Brynn, the researcher who handed over the dataset to me had taken an interest in the position of the inscriptions, so this was my next port of call. I first investigated what his

typical style was if he had one. We can observe that he had a preference to signing on the back of his works across his oeuvre and generally prefers to sign his paintings on the right hand side, rather than the left. It bears little importance to this project, but I would hazard to guess he was right-handed based upon this data. With particular focus on his habits during the 1940's we can see it resembles the period before it quite similarly.



I then mapped them over time, which we can observe here. The spike in signing on the back of the paintings in the sixties is quite fascinating. In the 1940s and 50s, we can see the steady increase in signing on the bottom right (rechtsonder) drops between 1935-1945 before shooting up again after.

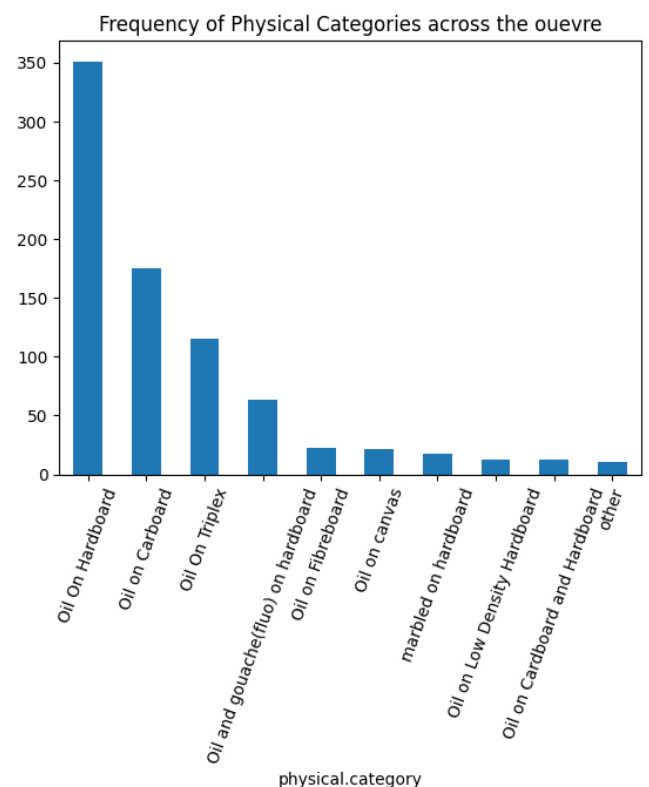
I next observed evolution of the dimensions of the paintings over time, using colour coding to track the temporal axes, and mapping the width on the x-axis and the height to the y-axis. This data is very strongly correlated as you would expect dimensions of 4 sided paintings. Yet even over time



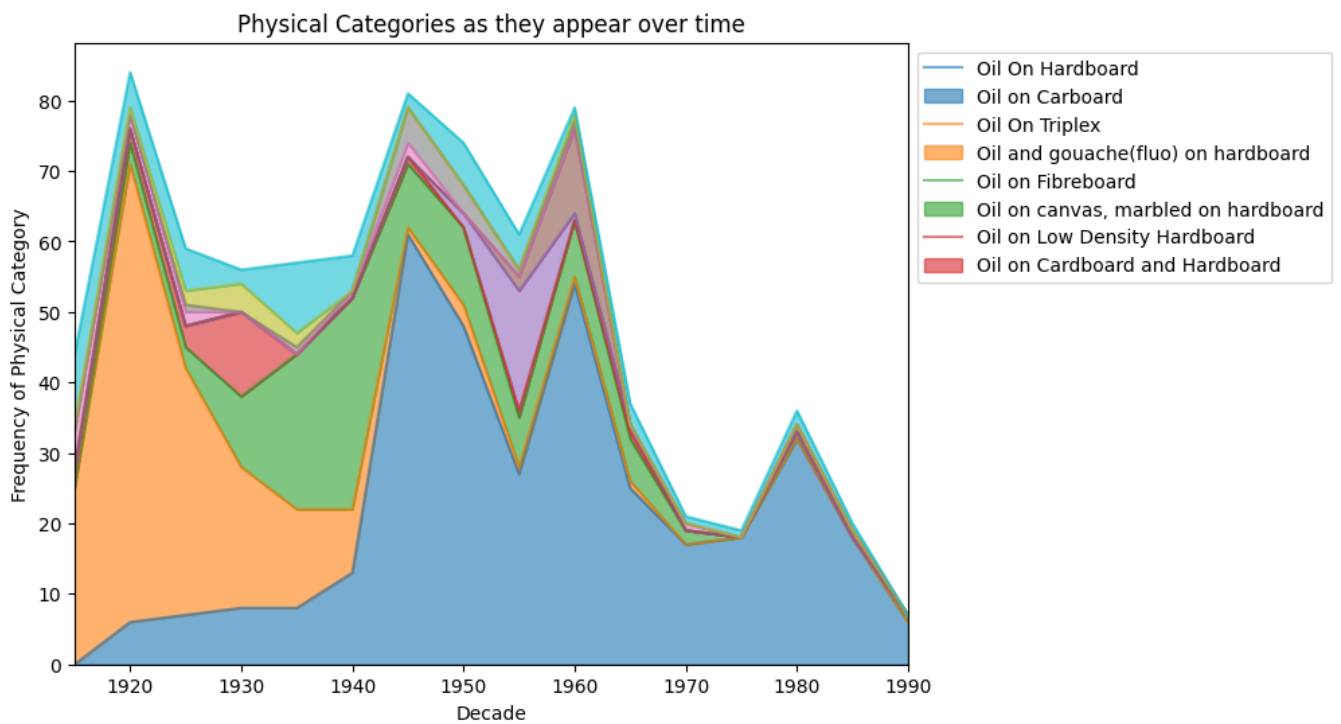
there is little that can be discerned as a trend. The only trend I see is that during his later years, shown in the yellow dots on these graphs he tended to have a preference for canvases that tended to be slightly less wide and a bit taller than you would expect. This may have been something he found easier to work on in his later years. We can also see that he may have been quite fond of canvases 60 cm wide, that or this was a standard size that was sold.

Aside from these findings, I do not see much that may help us ascertain if there is a correlation between the dimensions of the paintings and his potential backdating of paintings during our period of interest.

The following sets of graphs observes the different combinations of materials he used over the course of his oeuvre. We can see his preference for oils on full display, and the consistent use of hardboard.

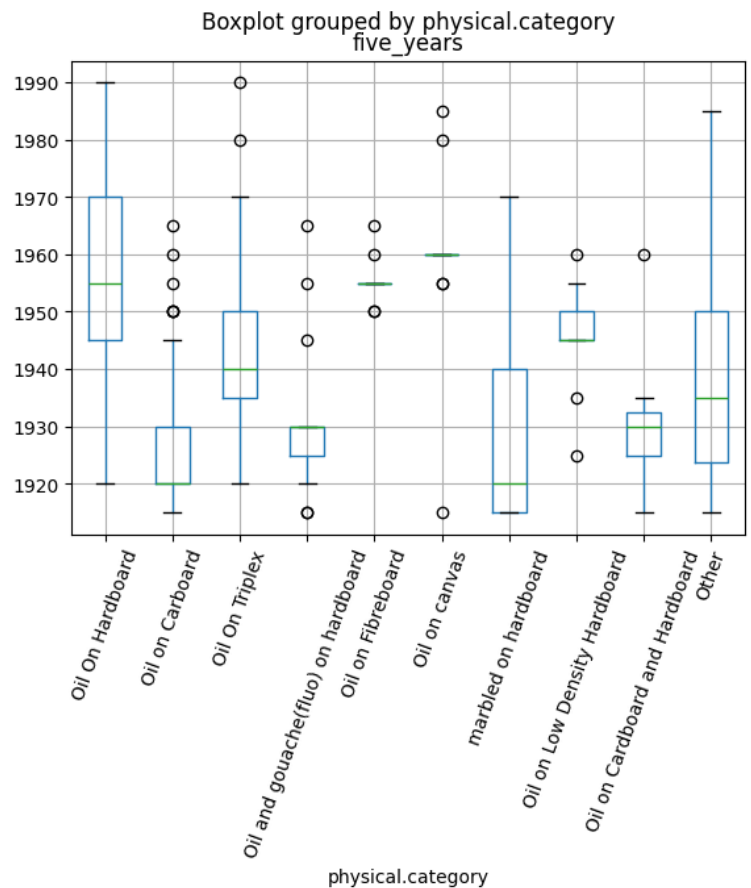






I then mapped all the physical categories together over time using a stacked area chart. Interestingly we can see that he does not start using his most used medium (oil on hardboard) with gusto until the forties, but it is still present in his earlier period. This may be due to lower supplies of fiberboard during this period, or perhaps he didn't like using it as much as he enjoyed using triplex. If I had access to sales records of the paintings from this period, I might be able to draw further conclusions, but one potential hypothesis is that these paintings on oil and hardboard from the period of 1915-1940 are likely to be backdated. I say this because due to supply issues/changes during the wartime period, acquiring triplex may not have been an option, or perhaps this is when hardboard became most available to him, which would explain the sharp increase and continued use after this time. If I had access to sales records, I could see that oil on hardboard paintings had been sold during this period and I could rule them out as being backdated, but without it I can only wonder.

This boxplot shows the distribution of each category over time. I had much higher hopes for this plot, but not much stands out aside from that one oil on canvas work that appears in 1915, since there is no other oil on canvas works for the next 5 decades.



Following this I went on to attempt to map the inscription categories, categorising each type of signature that appears across his oeuvre. Due to the issues I had outlined with copying and modifying the dataset at times, alongside the mismanagement of my project files I couldn't finish mapping all of the categories in time for the end of this project. I have created a number of graphs using this subset of the dataset, which I will put into the Appendix of this report. Included there is the beginning of a combinatorial look at both the physical categories and the inscription categories mapped together, still on this subset of the dataset.

## Reflections

To begin with I would again reiterate that the incompleteness of this research does devalue its merit. I feel that the graphs that combine the use of multiple variables such as physical categories and inscription categories would have been quite effective at shedding light on De Boeck's behaviours. The lack of a translator employed for aid on this project was also a large misstep; There was potential information in the dataset I may have missed, the majority of David Veltman's recent biography on Felix De Boeck was not usable, as well as

it leading to my not changing of a large amount of graph labels, making the graphs difficult to read for the English speaking readers this report is intended for. I do not regret avoiding the use of an AI translator for these items however, as they are unreliable. Although I admit, for the labels and other small translations I could have just taken a dictionary from the library, though overall I do not feel it detracts from the visualisations too much. One particularly irksome part of the visualisations is the colour labels for the legend on the stacked area charts in this report. Between changing the colour and tweaking the style of the legend I could not make it more clear what item on the area chart corresponded to the legend more clearly.

## Conclusion

I touched on some of the results drawn from the visualisations as I discussed them, they will be recapped here alongside other conclusions I have drawn.

Regarding the drop in painting activity during the period of 1940-1945 as shown in the first visualisation, I feel that this could be attributed to a lack of creativity during war-time among other events that were ongoing in his life during that time. Yet, according to Veltmans biography we know he was a man that does not balk at much (2021), and we can see from the same visualisation that he painted quite a lot as a younger man during war time. Due to the demand for art in Belgium as a way to spend money that might retain some value during this time as David and Oosterick (2012) pointed out, we may be able to understand this dip of painting during the period of 1940-1945 as simply a dip in truthful dating. This is by the very nature of this method of authentication, speculation.

There is one painting that used oil on canvas and is dated 1915, which I find particularly interesting as there was no other painting in his oeuvre using this technique for multiple decades after the fact.

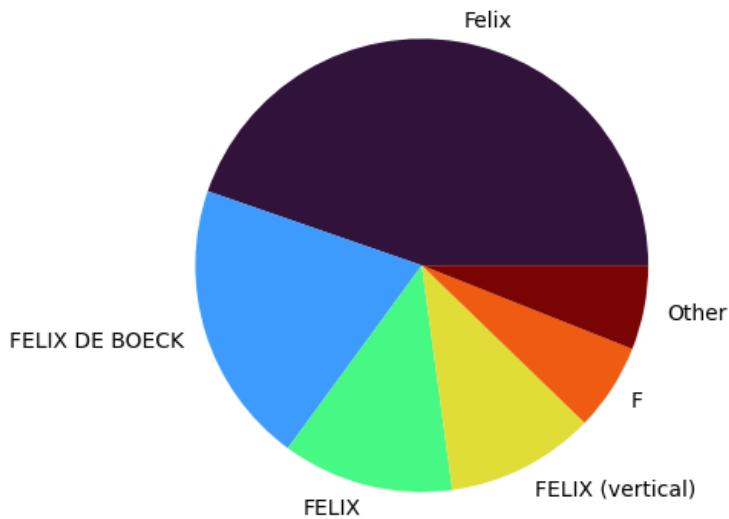
Finally, I would be interested in discovering what the sales records of oil on hardboard paintings are from the period of 1915-1940 to determine whether or not this subset of painting could potentially be backdated.

With regards to the development of a new technique with which to authenticate paintings, I feel as though I have proved that this method shows some promise. It can clearly show differences in behaviour over time. Of course, we must acknowledge that I specifically was hunting for potential examples of backdating, and when you ask a question, data will give meaning to it regardless of whether that data can hold your answer. The efficacy of this method will be discovered when this collection is fully authenticated, although I admit the lack of completeness in its study will hinder that progress. I feel as though the tools and techniques I selected did a good job in creating a replicable system of use, if someone wished to take this method and build upon it they would have this report as a resource as well as the code from my project to build upon. I stand by the decision to use openly accessible tools, though forcing the single streamline suite of tools is unnecessary in hindsight.

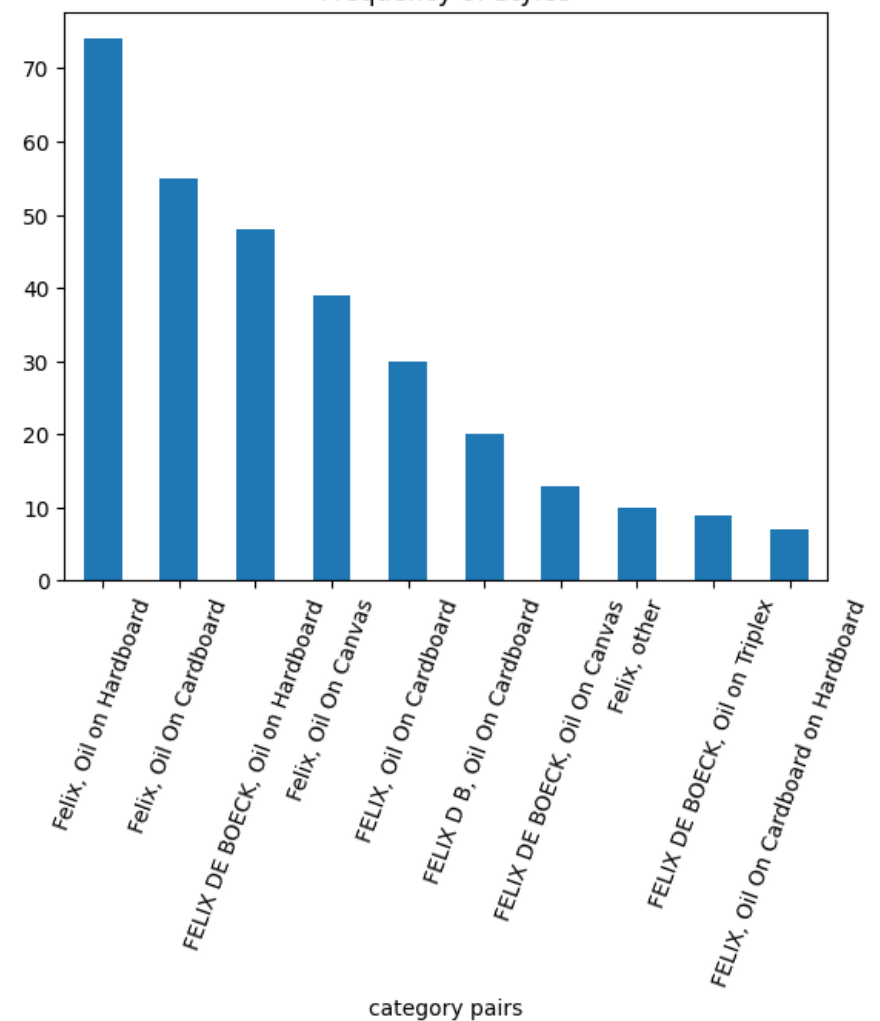
# Appendix

Graphs on incomplete dataset regarding physical categories and inscription categories.

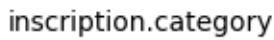
Relative Frequency of signatures across the oeuvre



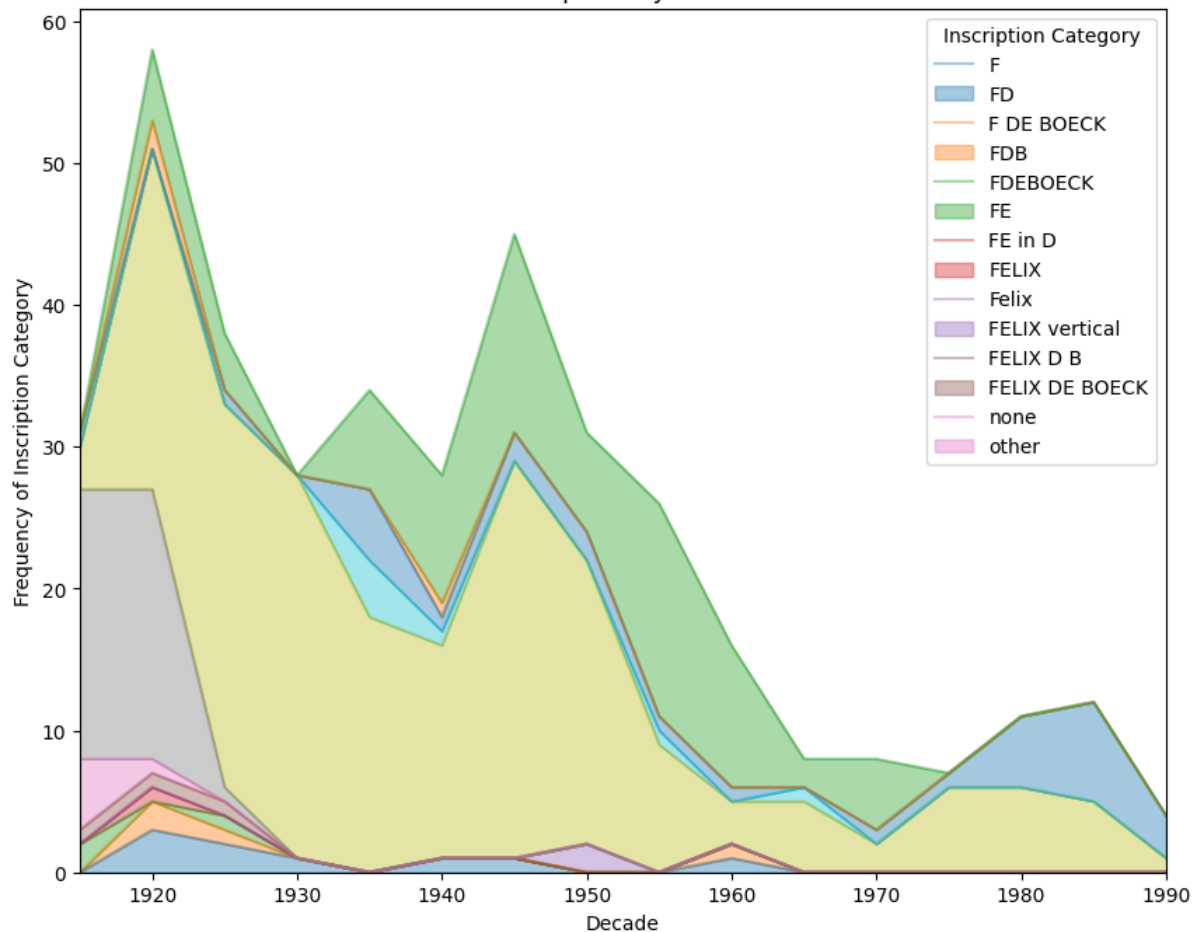
Frequency of Styles



five\_years



### Distribution of Inscription Styles across the Oeuvre



Google Drive link to full project base: <https://drive.google.com/drive/folders/1yIXypgroL6-dU82wZ5IKDPJxmW3QRKxX?usp=sharing>

## Reference list

David, G. and Oosterlinck, K. (2012). War, Inflation, Monetary Reform and the Art Market. *EHES Working Papers in Economic History*, [online] 12. Available at: <https://hdl.handle.net/10419/246943> [Accessed 20 Apr. 2025].

Guerra, M.F. (2008). Archaeometry and Museums: Fifty Years Of Curiosity and Wonder\*. *Archaeometry*, 50(6), pp.951–967. doi:<https://doi.org/10.1111/j.1475-4754.2008.00415.x>.

Polatkan, G., Jafarpour, S., Brasoveanu, A., Hughes, S. and Daubechies, I. (2009). Detection of forgery in paintings using supervised learning. In: *IEEE Xplore*. [online] pp.2921–2924. doi:<https://doi.org/10.1109/ICIP.2009.5413338>.

Robson, D. (1985). The Market for Abstract Expressionism: The Time Lag between Critical and Commercial Acceptance. *Archives of American Art Journal*, [online] 25(3), pp.19–23. doi:<https://doi.org/10.2307/1557387>.

Veltman, D. (2021). ‘*Sterven in het bed waarin ik geboren ben*’: Een biografie van Felix de Boeck (1898-1995). [Dissertation] *University of Groningen research database (University of Groningen / Centre for Information Technology)*. University of Groningen. doi:<https://doi.org/10.33612/diss.171646663>.

Windhager, F. (2020). *A Synoptic Visualization Framework for Artwork Collection Data and Artist Biographies*. [Dissertation] pp.106–130. Available at: <https://phaidra.univie.ac.at/detail/o:1393707.pdf> [Accessed 20 Apr. 2025].

Windhager, F., Federico, P., Schreder, G., Glinka, K., Dork, M., Miksch, S. and Mayr, E. (2019). Visualization of Cultural Heritage Collection Data: State of the Art and

Future Challenges. *IEEE Transactions on Visualization and Computer Graphics*, [online] 25(6), pp.2311–2330. doi:<https://doi.org/10.1109/tvcg.2018.2830759>.