Managing your Master thesis

dr. Hannes Datta
Tilburg School of Economics and Management
Department of Marketing
March 2018

TILBURG UNIVERSITY

Understanding Society

dr. Hannes Datta

Key areas of expertise
marketing mix modeling
causality in observational data
data management of structured and unstructured data
ownership versus access-based business models (streaming)

Educational background
PhD, quantitative marketing (Maastricht University)
Assistant professor Tilburg University (2013-...)

Teaching activities
MSc Marketing Analytics and Management: Data science skills for digital and social media research (2013-...)
BSc: Digital and social media strategies (2016-2017)

Research experience
Published in top tier journals *Journal of Marketing*, *Journal of Marketing Research*, *Marketing Science*
Most recently: online streaming (Spotify), see tiu.nu/spotify

Acknowledgements

This slide deck builds on the shoulders of giants. I am especially grateful to:

- Harald J. van Heerde, and his tutorial on “how to publish” (presentation at ANZMAC Doctoral Colloquium, 2010)
- Jacqueline van Beuningen (who came up with a PhD Survival Guide, in 2009)
- Thesis skills, developed by Jan Meyer (Tecnologica de Monterrey, Mexico) and Anne Klesse (Rotterdam School of Management, both 2010), and Johannes Bögershausen (developed in 2012)

The field of marketing

Key success factors

Content

finding or refining a topic, and carving out your contribution

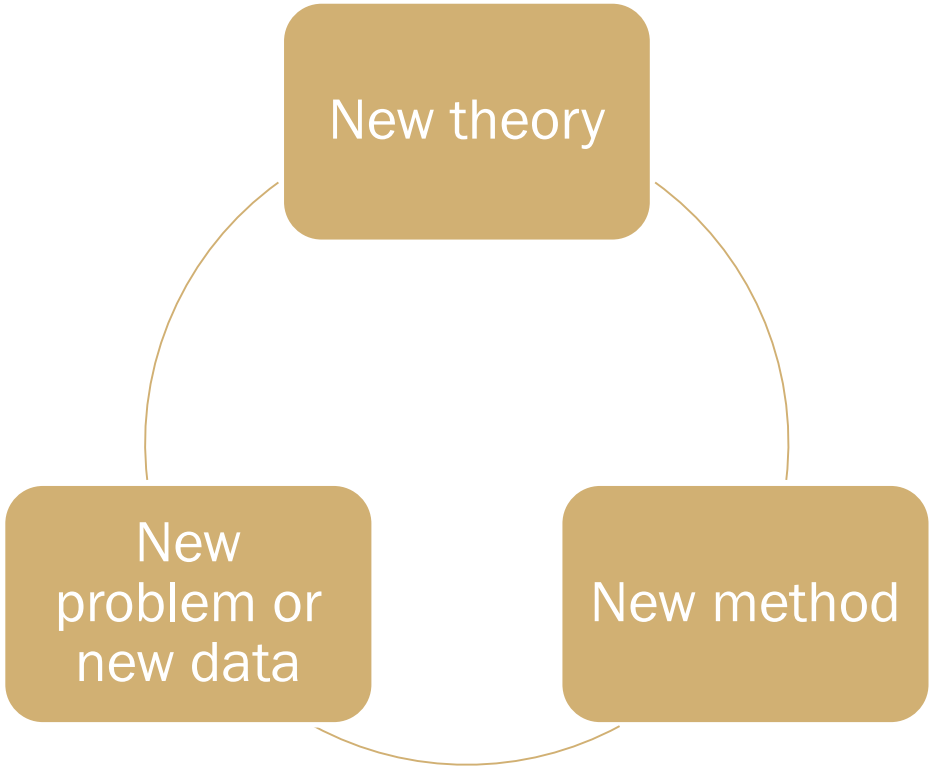
Finding/Refining your topic: Stage 1

- Find ideas…
  - In academic journals, in practitioner journals, at MSI (http://msi.org/), on the Internet (e.g., McKinsey, Accenture), the news, conversations, talk to managers, own experience, stuff mentioned in this presentation
- Talk to others about your ideas (“Do you find it interesting?”)
- Start writing ideas down
  - writing is not a linear process, first draft doesn’t have to be perfect!!!
- Realize that your ideas will continue to change (even after data collection)… and that’s good!
- Find your focus (narrow down your idea)
- Start thinking about processes and issues instead of topics (!)

Finding/Refining your topic: Stage 2

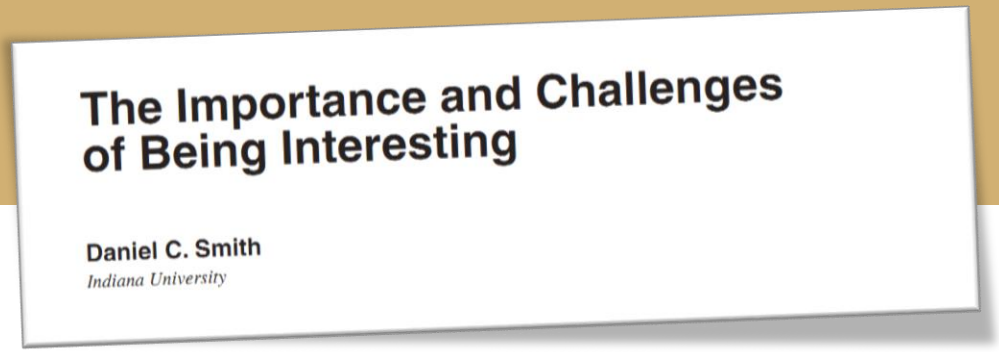
- Take control and ownership (this takes time!)
- Follow your heart
- Take your time to think
- Read (!!!) related and unrelated stuff you find interesting
- Try to structure your literature and take notes

Making a contribution (I)



"The important issue is not how <u>many</u> contributions a study will make, but rather the <u>significance</u> of each."

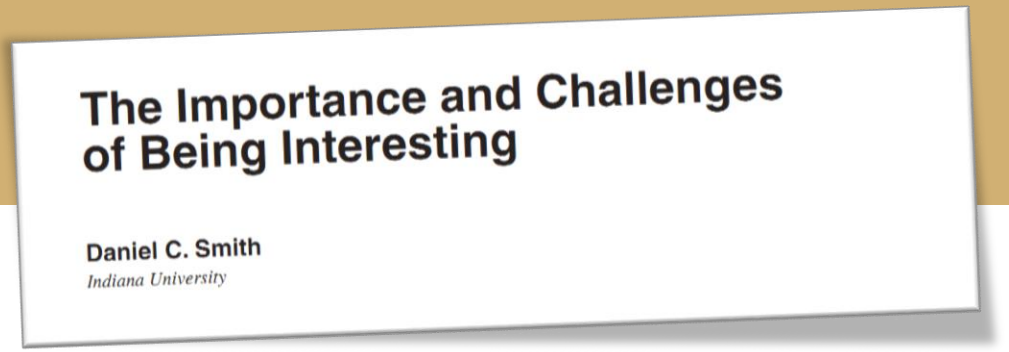
Making a contribution (II)


The Importance and Challenges of Being Interesting

Daniel C. Smith
Indiana University

- Test the assumptions
  … on which a significant stream of research relies
  - Taking a step back and identify core assumptions
- Probe the external validity
  - … of what we take to be true
  - Common background factors?
- The next new thing
  - Open new domains of inquiry
  - Make sure that this domain has significant implications if it was better understood

Making a contribution (III)


The Importance and Challenges of Being Interesting

Daniel C. Smith
Indiana University

- Work backward in the causal chain
  - Examine a variable, that if understood better, has major implications for a vast nest of relationships
- Intervene in an accepted causal chain
  - Taken-for-granted relationships? But really, really? What happens if they disappear?
- Challenging conventional managerial practices/beliefs
- Resolve inconsistent findings
  - Requires taking a step back…
  - Examine a stream of research in a holistic fashion
  - Focus on subtleties, that may have not so subtle implications

Do you have a contribution?!

a) Three questions-pitch

- Answer the following three questions about your research.
    1. What's new about it?
        - E.g., new data? new variables? new method?
    2. Is it useful (and to whom?)
    3. Is it interesting?
        - Simply telling something everyone knows already isn't interesting (known facts ≠ interesting)
        - Phenomenon, Gap, Aha!
        - Would you get your friends excited about it?!

b) Elevator pitch

- Imagine yourself going in the elevator and meeting the Rector of Tilburg University, who asks you to tell him what your MSc research is about…
- Have only 1-2 sentences, adhering to one of the following two structure:

  While …, our research shows that…

  Despite …, we find that…

Your pitch won't be perfect at the first time –
so practice it over and over again – until you really become enthusiastic about it!

outline of your thesis + chapter-specific tips

Outline of your thesis

1. Introduction
   - See next slides
2. Literature review
   - Elaborate discussion “what do we know”/”not know” and contribution (“what do you add to the literature”)
3. Conceptual framework
   - Boxes and arrows (for example, A leads to B, moderated by C)
   - Hypotheses or expectations
4. Data
   - Description of data collection
   - Variable operationalization
   - Descriptive statistics of the final data set
5. Model
   - Model expressed in formulas + explanation
   - Explanation of variables that enter your model, and why
6. Results
   - See next slides
7. Discussion
   - Summary of main findings
   - Theoretical and managerial takeaways
   - Limitations and future research

This format may slightly deviate from the traditional thesis format, but is preferred with quantitative theses.

Chapter 1: Introduction

- How to write it up… (e.g., in your introduction or proposal)

1. Practical phenomenon / problem
   - Establish the importance of the area of interest and the phenomenon
   - Mention your research question indirectly (see grey box below)
2. Why to study it?
   - Motivate your research question
   - Start with something like… “Studying [xyz] is crucial to […]” or “The […] is worth studying because…”
3. What we know
   - Indicate in general terms what has been done in this area
   - Start with something like… “My research relates to extant literature… (in at least three ways): First, […]” / “My research contributes to two literature streams:”…
4. What we don’t know
   - Identify important gaps, inconsistencies, and/or controversies in the relevant literature
   - Provide a concise statement of the manuscript’s purpose(s) and the contributions the manuscript makes to the literature
   - Start with something like… “Our research extends extant research by…” / “Therefore, as a first contribution, we […]”
5. What you will do
   - Describe (a) which data you will use, (b) which methods you will use to study your research question. Then, (c), describe the flow of the next sections.

- Do not use subheadings in your introduction, but structure it in **paragraphs**
- **Do not explicitly formulate your problem statement** (e.g., “therefore, the following problem statement is formulated: *“Which drivers affect […]”*). Rather, define your problem statement *indirectly*: “Therefore, it is crucial to study the drivers affecting […].”. This is much more natural.

Chapter 2: Literature review

Wrong approach

- Study A finds 1
- Study B finds 2
- Study C finds 1
- Study D finds 1+2
- Study E finds 3
- Study F finds 4
- My study contributes to the literature by doing x, y, and z.

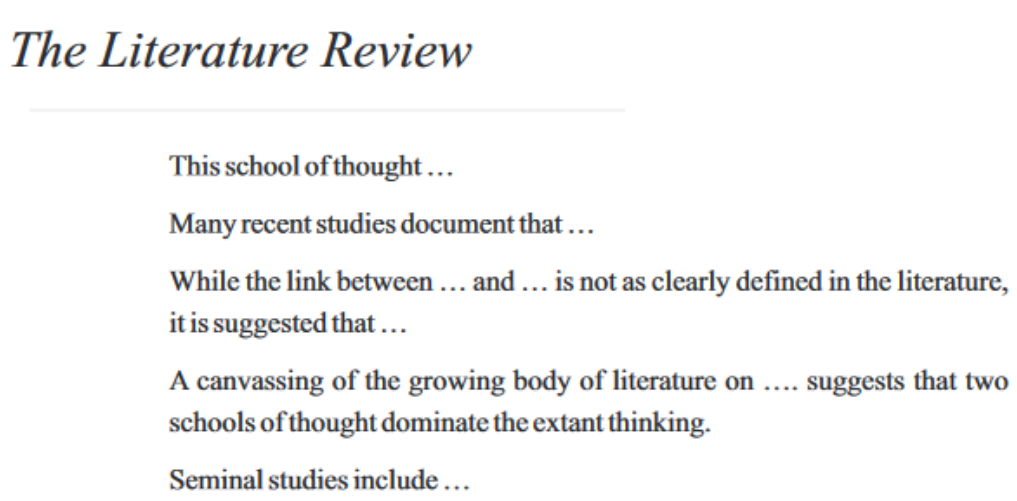
Correct approach

- First stream of literature is concerned with 1.
  - Study A supports 1 (+ argument used in study)
  - Study C adds to it (+ argument used in study), providing further evidence for 1.
  - **But,** they miss out on x (→ contribution!)
- Second stream of literature is concerned with 2.
  - For example, study B and D find 2, while study D even supports 1.
  - **But,** they miss out on y (→ contribution!)
- Third stream deals with related phenomena.
  - Study E finds 3 (+ why it's important to your study)
  - Study F finds 4 (+ why it's important to your study)
  - **But,** they miss out on z (→ final contribution!)

Key is to synthesize the literature!

Chapter 2: Literature review

- Most students I talk to enjoy writing their literature review section least.
- So, make it a little more fun by using a few "stock phrases" – this will help you to submit a polished draft as opposed to a lousy set of basic sentences.
- As a preview, see this:



The Literature Review

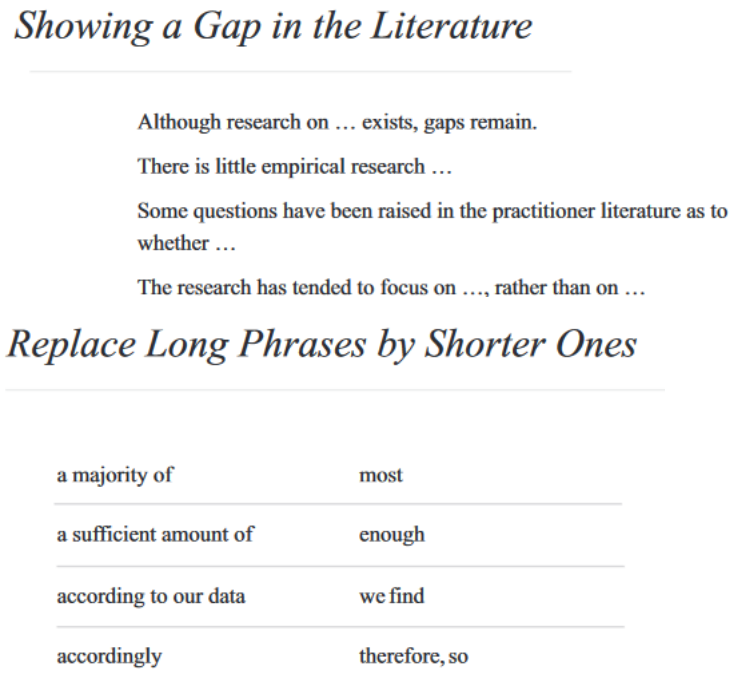
This school of thought …

Many recent studies document that …

While the link between … and … is not as clearly defined in the literature, it is suggested that …

A canvassing of the growing body of literature on …. suggests that two schools of thought dominate the extant thinking.

Seminal studies include …



Showing a Gap in the Literature

Although research on … exists, gaps remain.

There is little empirical research …

Some questions have been raised in the practitioner literature as to whether …

The research has tended to focus on …, rather than on …

Replace Long Phrases by Shorter Ones

a majority of	most
a sufficient amount of	enough
according to our data	we find
accordingly	therefore, so

- More to be found here: tiu.nu/writing

Chapter 2: Visualizing your contribution (I)

Table X:

CLASSIFYING PAST CLV STUDIES

	Products	**Services**
No distinction free trial vs. regular	Study 1 Study 2	Study 3
Distinction free trial vs. regular	Study 4 Study 5 Study 6	**This study**

Chapter 2: Visualizing your contribution (II)

Table 1: Current Literature on Long-term Effects of Marketing Variables

	Effect of				Effect on	Modeling Approach	# Categories
	Promotion	Advertising	Distribution	Product			
Clarke (1976)		✓			Brand Sales	VPM	1
Baghestani (1991)		✓			Brand Sales	VAR	1
Dekimpe and Hanssens (1995)		✓			Chain Sales	VAR	1
Papatla and Krishnamurthi (1996)	✓				Choice	VPM	1
Mela, Gupta, and Lehmann (1997)	✓	✓			Choice	VPM	1
Mela, Jedidi, and Bowman (1998)	✓				Incidence and Quantity	VPM	1
Mela, Gupta, and Jedidi (1998)	✓	✓			Market Structure	Mixed	1
Kopalle, Mela, and Marsh (1999)	✓				Brand Sales	VPM	1
Jedidi, Mela, and Gupta (1999)	✓	✓			Choice and Quantity	VPM	1
Foekens, Leeflang, and Wittink (1999)	✓				Brand Sales	VPM	1
Dekimpe and Hanssens (1999)	✓	✓			Brand Sales	VAR	1
Dekimpe, Hanssens, and Silva-Risso (1999)	✓				Brand and Cat. Sales	VAR	4
Srinivasan, Leszczyc, and Bass (2000)	✓		✓		Market Share	VAR	2
Bronnenberg, Mahajan, and Vanhonacker (2000)	✓	✓	✓		Market Share	VAR	1
Nijs et al. (2001)	✓				Category Sales	VAR	560
Pauwels, Hanssens, and Siddarth (2002)	✓				Incidence, Choice and Quantity	VAR	2
Srinivasan et al. (2004)	✓				Margin and Revenue	VAR	21
Pauwels (2004)	✓	✓		✓	Brand Sales	VAR	1
Van Heerde, Mela, and Manchanda (2004)				✓	Market Structure	VPM (DLM)	1
Pauwels et al. (2004)	✓			✓	Financial measures	VAR	1
Steenkamp et al. (2005)	✓	✓			Brand Sales	VAR	442
Sriram, Balachander, and Kalwani (2007)	✓	✓		✓	Brand Sales	VPM	2
Ataman, Mela, and Van Heerde (2008)	✓	✓	✓	✓	Brand sales (new brands only)	VPM-SE (DLM)	22
Slotegraaf and Pauwels (2008)	✓			✓	Brand Sales	VAR	7
THIS PAPER	✓	✓	✓	✓	Brand sales and Elasticity	VPM-SE (DLM)	25

Notes: VPM = Varying Parameter Model; VAR = Vector Autoregressive model; DLM = Dynamic Linear Model; SE = System of Equations

Ataman, Berk, Harald J. van Heerde, Carl. F. Mela (2010), "The Long-term Effect of Marketing Strategy on Brand Sales,", Journal of Marketing Research.

(Please not that your table may have different column headings than this one!)

Chapter 3: Conceptual framework

- Have theory of why you expect a relationship between your Xs and Y(s)
  - theory can be based on papers, but also your own logical reasoning, as long as it motivates the relationship
  - also think about alternative arguments (X could have a positive relationship, but also a negative relationship)
  - write up your argumentation, then, end with your hypothesis, or spell out your expectation explicitly (you can also admit the relationship can go in two ways)
- Typically, just hypothesizing an effect to exist is weak; it's much stronger if you can predict its direction
- Draw a conceptual framework (and re-draw until it's perfect)

Hypothesis format

- If you have a strong theory, you can cleanly predict what happens to Y if you change X.
- In that case, it makes most sense having formal hypotheses in your paper.
- For example, see Datta, Foubert and van Heerde (2015, JMR, https://doi.org/10.1509/jmr.12.0160)

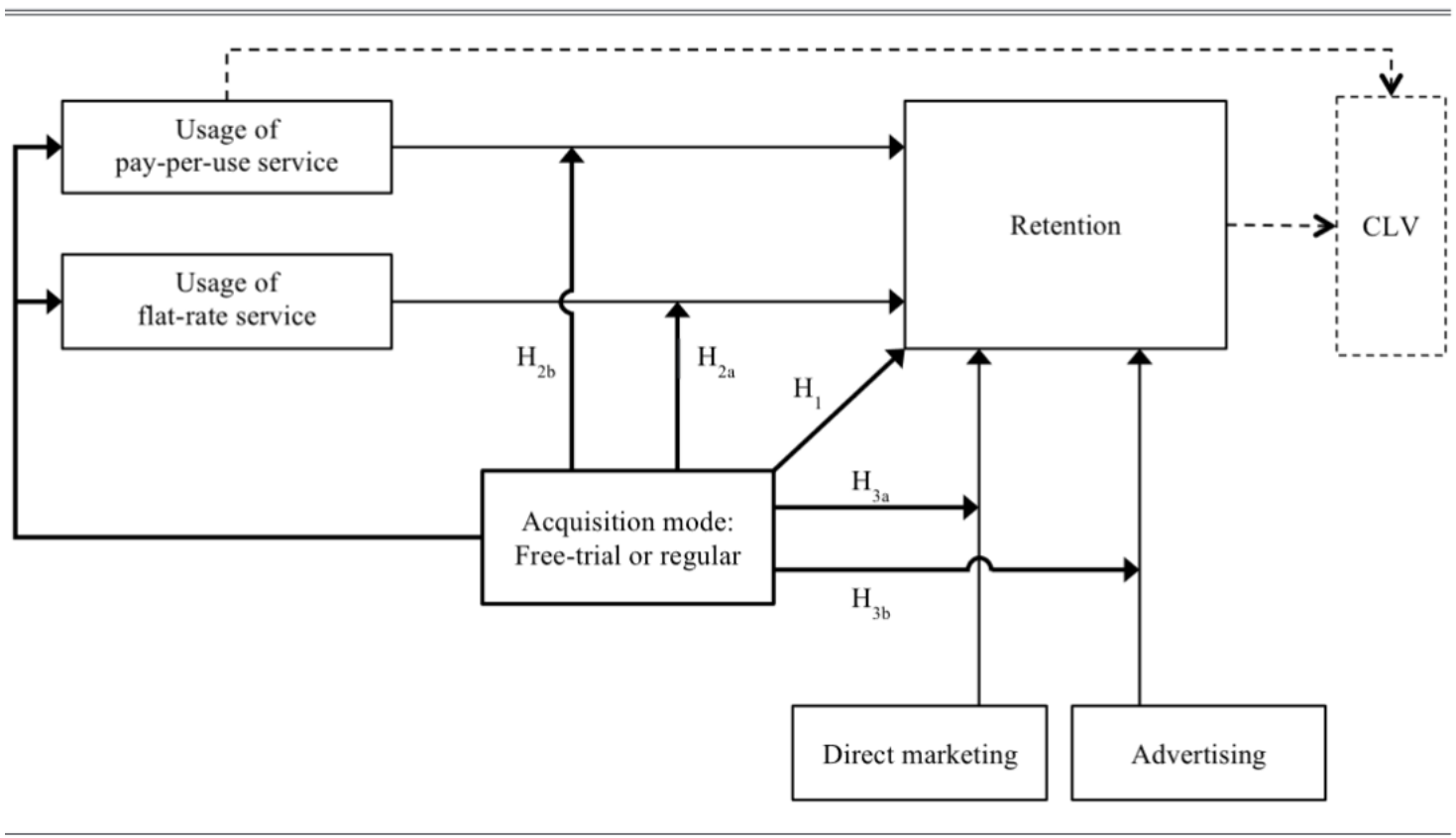
Expectations

- If your theory is not strong, or your predictions are frequently going in both ways (e.g., both positive and negative), it's wiser to not write out formal hypotheses but rather expectations.
- For a paper having expectations, see Datta, Ailawadi, and Van Heerde (2017, paper here). Alternatively, a paper can very well do without any hypotheses or expectations, see tiu.nu/spotify

choose a format that fits your research

Chapter 3: Conceptual framework (Example 1/2)

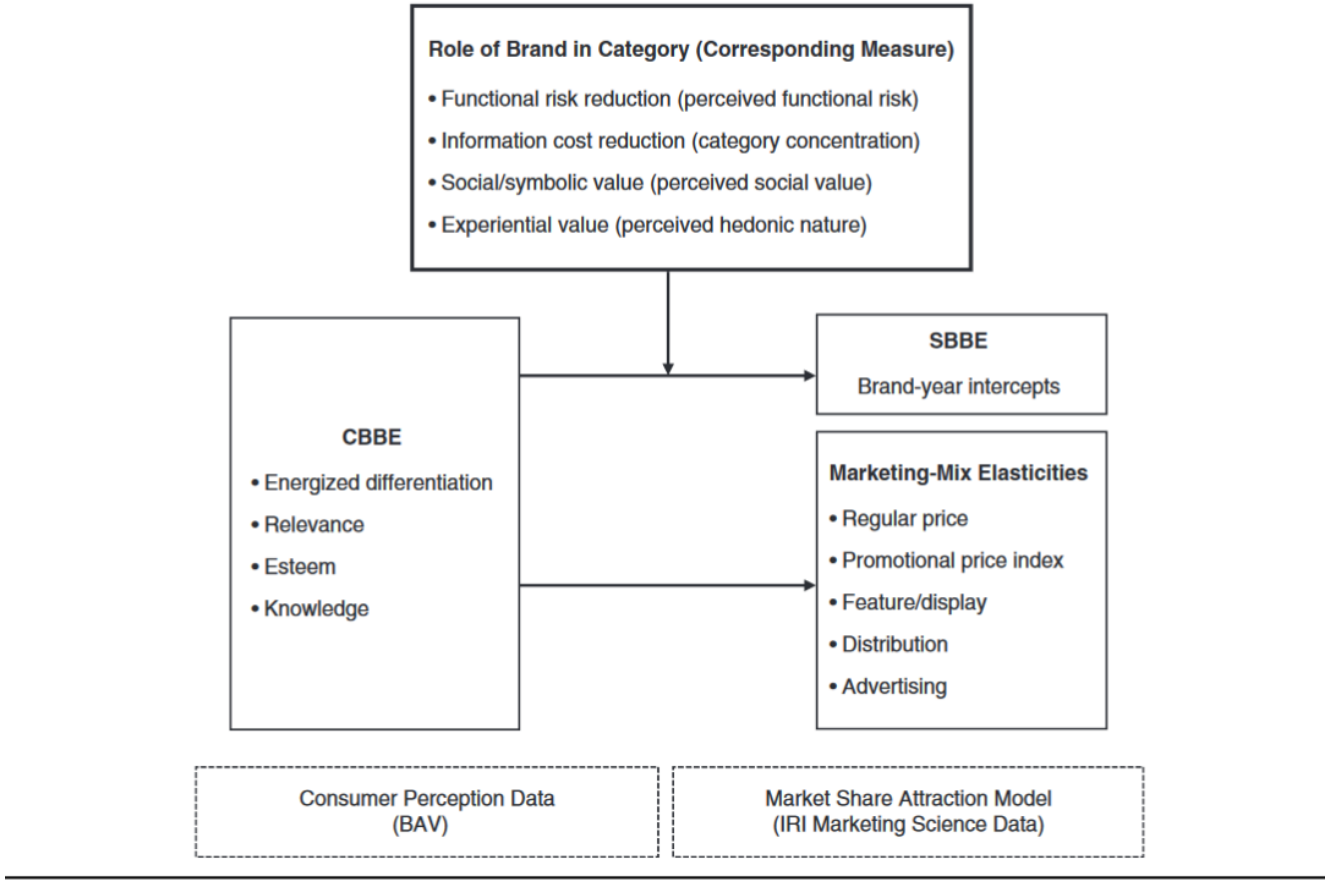
Figure 1
CONCEPTUAL FRAMEWORK



Datta, H., Foubert, B., & Van Heerde, H. J. (2015). The challenge of retaining customers acquired with free trials. Journal of Marketing Research, 52(2), 217-234.

Chapter 3: Conceptual framework (Example 2/2)

FIGURE 1
Guiding Framework



Hannes Datta, Kusum L. Ailawadi, and Harald J. van Heerde (2017) How Well Does Consumer-Based Brand Equity Align with Sales-Based Brand Equity and Marketing-Mix Response?. Journal of Marketing: May 2017, Vol. 81, No. 3, pp. 1-20.

Chapter 4: Data

1. Description of data collection
   - Describe how the data was collected (e.g., web scraping, APIs, a cooperation with the company)
   - Describe which data was collected ("the raw data")
     - E.g., time frame, number of cross-sectional units, type of data
   - As an example, see how I describe my own data collections in published papers

2. Data preparation and variable operationalization
   - Describe how you move from the raw data set to the final data set (see next two slides for an explanation)

3. Provide descriptive statistics of the final data set

Please have a look at Gordon, Brett R., Avi Goldfarb, and Yang Li. "Does price elasticity vary with economic growth? A cross-category analysis." Journal of Marketing Research 50.1 (2013): 4-23. This is a very well written description of how the data was prepared. Take it as an example/motivation of how you could write up your data section, too.

Tips: Describe your raw data

- It is crucial that the reader (and your supervisor) understands the format of your data.
- You need to distinguish between your raw data, and your final data set.
- Your raw data
  - is how the data is stored at the company, or how you gathered the data yourself (e.g., using web scraping technology)

  You need to know:
  - What's the "primary key" of this data? (→ what identifies a unique row in this data set?)
    - For example, data may be stored per video_id – day (e.g., the number of YouTube views per video per day), or per shop – user – day (recording sales of a user for an online shop per day)
    - Make explicit the frequency of your data (e.g., per month, week, day, hour, second…)
  - What are the "value" columns?
    - Value columns is data that is recorded per primary key (e.g., video views for YouTube, sales for the online shop).
  - Typically, you may encounter different tables with different primary keys and value columns
    - E.g., a table with user demographics, a table with sales data, a table with clickstream data, etc.

- Create some summary statistics of this data
  - Always deliver
    - A table of mean, SD, min, max per variable ("descriptive statistics")
  - Try to be creative
    - E.g., for sales data of a shop, create a summary of how many users buy per shop, or
    - E.g., for panel data (i.e., users/brands/artists observed over time): some time-series plots (e.g., line graphs, for each user a line over time)

Tips: Preparing your final data set

- You will use a statistical program (e.g., SPSS, R, STATA) to transform your raw data to your final data set that you will analyze
  - Note, in many settings, it really does not make sense to analyze the raw data; it is (a) way too fine-grained (e.g., recorded per minute, while a weekly level of analysis may be warranted), (b) way too messy (e.g., contains outliers), and (c) does not yet contain the independent variables you are actually interested in
- Tasks to arrive at your final data set
  - Data cleaning
    - E.g., refine your sample (e.g., "shops that sell at least 1 item per week"; "brands that are in the Top 3 for at least three consecutive years", etc.
    - E.g., define rules how to deal with missing values and outliers
  - Data aggregation
    - E.g., aggregate from minutes to weeks; aggregate from sales per user per shop, to sales per shop ("same primary key")
  - Data merging
    - E.g., merge different sources (which have previously been aggregated to the same primary key); e.g., temperature data (recorded per day), to your sales data set for swimming equipment ;) (per shop, and day)
  - Operationalize your variables
    - Which are the variables you want to use for your analyze, and how do you operationalize them? (e.g., think of raw data that stores the names of products sold in a given month → you could convert this to a measure of *how many products* are sold in a given month, simply by counting them. So from your raw data, you get to a real variable that you can use (e.g., note the previous variable was text, and now you have a count, e.g., number of SKUs)
    - Typically, you provide a table with variable names, and your operationalization.
    - Look at the literature and how previous researchers have defined variables that you are looking for.

Chapter 4: Variable operationalization (Example 1)

Table 1. Variable Operationalization

Dimension	Operationalization
(1) *Quantity*	• Log number of song plays
(2) *Variety*	
Breadth	• Log number of unique artists, songs, and genres listened to
Concentration in common favorites (superstars)	• Number of unique artists in the top 20, top 100, and top 500 in a user's geographic region[a] (ranked according to a rolling window of a year, lagged by four weeks; $t-55,\ldots,t-4$), divided by the number of unique artists listened to over the same time period
Concentration in personal favorites	• Herfindahl index (sum of squared listening shares), computed over a user's weekly plays of artists, songs, and genres
(3) *Discovery*	
New content consumption	• Number of distinct new artists, songs, and genres listened to by a user for the first time,[b] divided by the total number of distinct artists, songs, and genres listened to
Repeat consumption	• Number of unique new artists, songs, and genres played more than once, divided by the total number of unique new artists, songs, and genres listened to
Best discoveries	• Amount of plays of the top 1 new artist, song, and genre in an 8-week period subsequent to discovery ($t+1,\ldots,t+8$) ranked in order of plays, divided by the amount of plays of the overall (not necessarily new) top 1 artist, song, and genre over the same time period

Note. All variables are computed at the user-week level.
[a]Regions are the European Union, South America, United States and Canada, and others.
[b]This refers to content consumed in a user's first week of consumption, based on the users' listening history on the service up to January 6, 2013.

Datta, H., Knox, G., & Bronnenberg, B. J. (2017). Changing their tune: How consumers' adoption of online streaming affects music consumption and discovery. Marketing Science.

Chapter 4: Descriptive statistics (Example 1/2)

TABLE 2
Sample Description

Category	No. of Brands	No. of BAV Brands	Mean No. of Years per Brand	Mean of Social Value[a]	Mean of Hedonic Nature[a]	Mean of Functional Risk[a]	Mean of Cat. Concentration[a]
Beer	59	37	9.9	3.39	5.96	3.44	.47
Carbonated Soft Drinks	27	21	9.8	2.72	5.32	3.01	.56
Cigarettes	25	21	10.0	3.11	4.26	3.10	.65
Coffee	30	23	9.4	3.07	5.52	3.64	.74
Cold (RTE) Cereal	23	20	10.0	2.72	4.75	3.16	.48
Deodorants	19	17	9.7	2.44	2.83	3.53	.51
Disposable Diapers	6	4	9.2	2.45	2.10	3.72	.99
Household Cleaners	15	9	9.7	2.59	2.18	3.41	.59
Ketchup	5	3	10.0	2.05	3.45	2.58	1.00
Laundry Detergents	20	17	9.9	2.56	2.40	3.53	.60
Margarine & Spreads	13	6	10.0	2.17	3.09	2.72	.65
Mayonnaise	7	3	10.0	2.15	3.07	2.83	.93
Milk	19	4	9.9	2.34	3.39	2.81	.90
Mustard	12	5	10.0	2.12	3.30	2.46	.90
Peanut Butter	11	5	9.4	2.30	4.37	2.98	.92
Frozen Pizza & Dinners	26	15	9.4	2.47	3.61	3.22	.47
Razors & Blades	5	3	10.0	2.52	2.61	3.72	.99
Salty Snacks	17	7	9.6	2.56	5.16	2.97	.74
Shampoo	28	17	9.1	2.84	3.09	3.56	.57
Soup	8	6	9.0	2.16	3.40	2.94	.98
Pasta Sauce	15	14	9.6	2.30	3.89	3.05	.70
Sugar Substitutes	10	5	8.1	2.52	2.82	2.76	.89
Toilet Tissue	10	5	10.0	2.45	2.40	3.60	.70
Toothpaste	15	13	9.9	2.53	2.79	3.43	.87
Yogurt	16	10	8.9	2.43	4.03	3.01	.79

[a]Category concentration is the total market share of the top four brands in a category. Both social value and functional risk of a category are measured on two-item, five-point Likert scales (1–5), with higher values representing higher scores. Hedonic nature is measured on a two-item, seven-point sematic differential scale (1–7), with higher values representing more hedonic categories. For measurement details, see Web Appendix B.

Hannes Datta, Kusum L. Ailawadi, and Harald J. van Heerde (2017) How Well Does Consumer-Based Brand Equity Align with Sales-Based Brand Equity and Marketing-Mix Response?. Journal of Marketing: May 2017, Vol. 81, No. 3, pp. 1-20.

Chapter 4: Descriptive statistics (Example 2/2)

Table 2

DESCRIPTIVE STATISTICS FOR FREE-TRIAL AND REGULAR CUSTOMERS

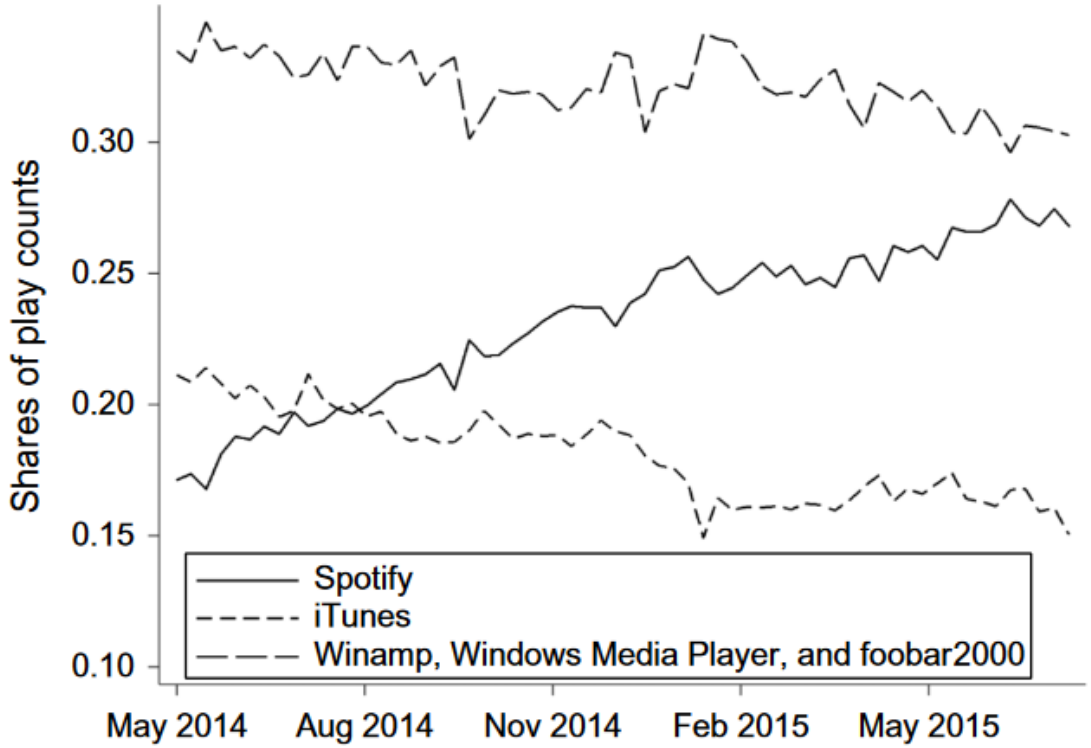
	Free-Trial Customers		*Regular Customers*	
	M	*SD*	*M*	*SD*
Dependent Variables				
Retention probability	.93	.25	.96	.19
Usage of flat-rate service (channel zaps)	169.35	176.90	189.28	183.80
Usage of pay-per-use service (number of VODs)	.73	2.19	.58	1.91
Marketing Communication				
Direct marketing (number of direct-marketing contacts)	.34	.55	.16	.41
Advertising (share of voice)	.74	.28	.79	.26
Customer-Specific Variables				
Age (years)	46.22	12.64	44.89	12.60
Household size	2.98	1.48	2.88	1.53
Annual income (in €10,000)	2.44	.56	2.43	.59
Time to adoption (in months following the launch of the service)	13.14	2.45	12.38	2.20
Control Variables				
Monthly subscription fee (in €)	9.66	7.80	13.22	4.85
Cancellation penalty (termination fees – future fees, in €)	–59.80	68.08	–8.14	25.02
VOD credit (in €)	5.50	8.90	1.38	5.08
Temperature (in degrees Celsius)	12.82	4.97	12.83	4.99
Time since adoption (in months)	5.34	3.52	5.95	3.68
Direct-marketing contacts before acquisition	.62	.61	.26	.48
Advertising (share of voice) before acquisition	.73	.27	.76	.27
Total fees for regular 12-month contract at the time of acquisition (in €)	193.83	50.04	181.97	44.79
Number of customers	12,612		3,900	

Datta, H., Foubert, B., & Van Heerde, H. J. (2015). The challenge of retaining customers acquired with free trials. Journal of Marketing Research, 52(2), 217-234.

Chapter 4: Visualizations

- Also consider plotting interesting aspects of your data, e.g., my paper on Spotify

Figure 1. Market Shares (Play Counts) in the Raw Data



Datta, H., Knox, G., & Bronnenberg, B. J. (2017). Changing their tune: How consumers' adoption of online streaming affects music consumption and discovery. Marketing Science.

Chapter 5: Method

- Typically, you will describe your model in a formula, and an accompanying text, e.g.:

$$(1) \qquad MS_{bt} = \frac{A_{bt}}{\sum_{j=1}^{M} A_{jt}}, \text{ and}$$

$$(2) \quad A_{bt} = \exp\left(\sum_{y \in Y_b} \alpha_{by} \times DumYear_{ty} + \beta_{b1} RegPrice_{bt} + \beta_{b2} PriceIndex_{bt} + \beta_{b3} FD_{bt} + \beta_{b4} Distr_{bt} + \beta_{b5} AdStock_{bt} + \sum_{a,l} \gamma_{al} Attr_{bal} + \sum_{q=2}^{4} \kappa_{bq-1} Quarter_{qt} + \sum_{k} \omega_{kb} Copula_{kbt} + \varepsilon_{bt} \right),$$

where we drop the category index c to simplify exposition, and

MS_{bt} = Unit market share of brand b in week t;

A_{bt} = Attraction of brand b in week t;

α_{by} = Brand- and year-specific intercept for brand b in year y;

$DumYear_{ty}$ = Indicator variable, equal to 1 if week t is part of year y, and 0 otherwise;

$RegPrice_{bt}$ = Regular price of brand b in week t, deflated by a category-specific Consumer Price Index to account for category-wide price changes;

$PriceIndex_{bt}$ = Actual price of brand b in week t divided by its regular price to measure its promotional price discount;

FD_{bt} = Intensity of feature and/or display support for brand b in week t;

$Distr_{bt}$ = Total distribution of the stockkeeping units (SKUs) of brand b in week t;

$AdStock_{bt}$ = Smoothed advertising spending or advertising stock of brand b in week t, where $AdStock_{bt} = \lambda AdStock_{b,t-1} + (1 - \lambda)Advertising_{bt}$;

$Attr_{bal}$ = Fraction of the SKUs of brand b that have attribute level l for attribute a;

$Quarter_{qt}$ = Quarterly dummy for quarter (q = 1 if week t is in quarter q, and 0 otherwise), mean-centered at the brand level;

$Copula_{kbt}$ = Gaussian copula (control function term) for marketing-mix variable k of brand b in week t to control for potential endogeneity of the variable; and

ε_{bt} = Normally distributed error term for brand b in week t.

Source: Hannes Datta, Kusum L. Ailawadi, and Harald J. van Heerde (2017) How Well Does Consumer-Based Brand Equity Align with Sales-Based Brand Equity and Marketing-Mix Response?. Journal of Marketing: May 2017, Vol. 81, No. 3, pp. 1-20.

- Pay attention to use the correct sub indexes (e.g., i for consumers, j for categories, t for time, etc.) – not everything will be ijt!!!

Chapter 6: Results

1. Model fit
   - E.g., R2, Adj. R2, F-test (for regressions)
   - E.g., Log-likelihood test against a null model, hit rate, hit probabilities (for Logit models)

2. Reporting your results
   - Have a result table. Typically, this is not the output you get from SPSS or R, but a table you compile yourself.
     - E.g., a table combining multiple models
     - E.g., a table dropping those 100 controls you added but that are not worth discussing
     - E.g., adding fit metrics to the table
   - It's important to consider the exact situation you are in, and making the table fit your purpose

Chapter 6: Results

3. Explaining your results
   - Hypothesis tests (do this for each hypothesis) / and or revealing results about your expectation
     - Briefly repeat hypothesis ("In H1, we hypothesized that [...]")
     - Report the result in brackets (e.g., "Indeed, the effect of A on B is statistically significant (beta = xx, *p* = .012)")
     - Explain the result (the "why?")
       - For a confirmed hypothesis, increase the intuition of your hypothesis ("apparently, as hypothesized, A leads to B **because** [repeat, etc.]...")
       - For an unconfirmed hypothesis, explain why you don't find the effect (e.g., a conceptual reason (effect doesn't exist, and you have an argument why it may not exist), or a measurement issue (e.g., data problems), etc.)

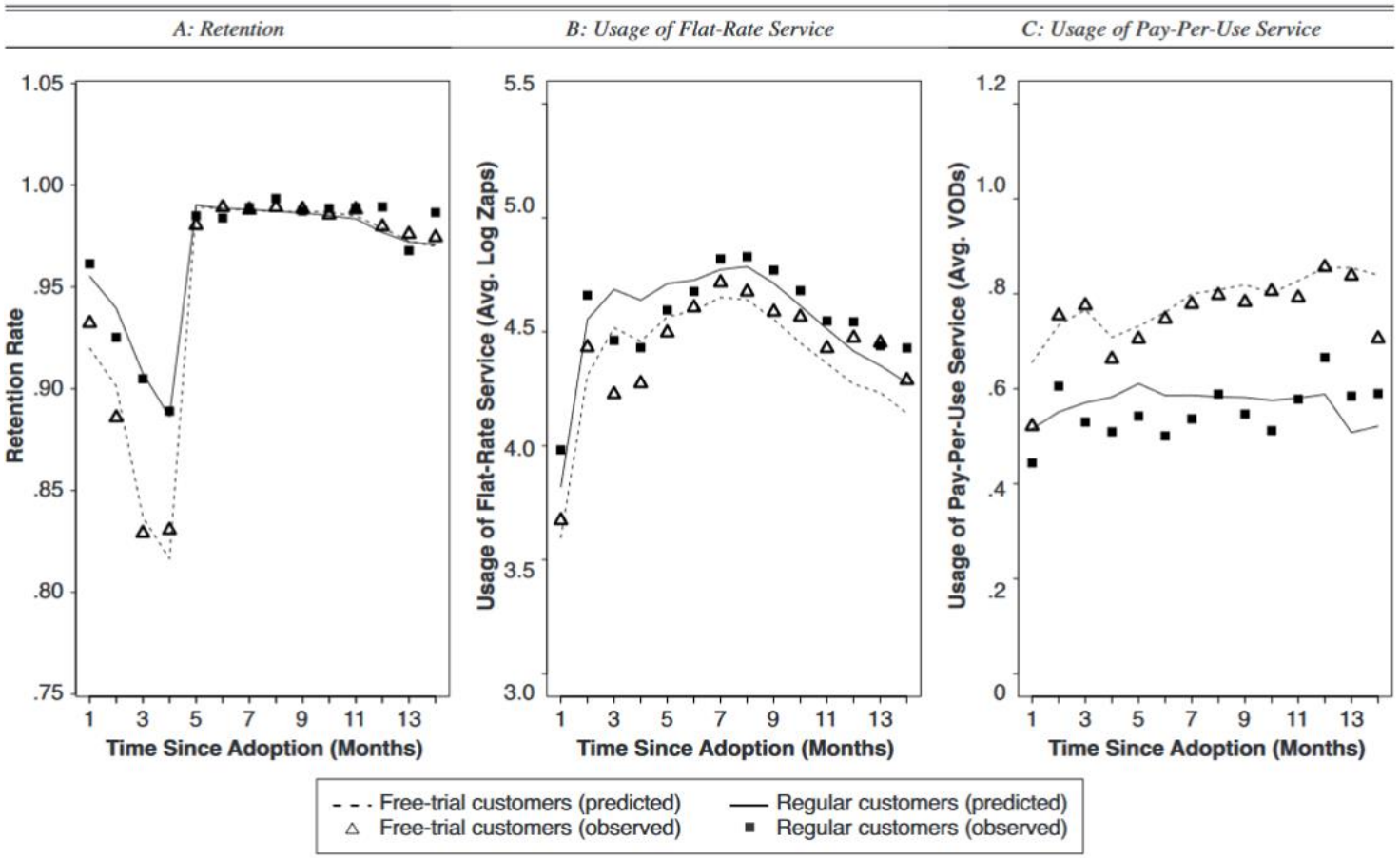
   - Remaining variables
     - Explain the effects of control variables here ("The control variables age and gender have face-valid effects. For example, all else equal, age increases the intention to purchase (beta = xx, *p* = .12). Education turns out to be not a significant predictor of intention to purchase (beta = xx, *p* = .63). This may be the case because [...]

- Note: While hypothesis tests are usually carried out against a *p*-value of .05 (or sometimes .10), you need to report the *exact* p-value in the text and tables with, e.g., three digits after the decimal point (e.g., *p* = .049, instead of *p* < .05). If you cannot give an exact *p* value because of rounding (e.g., the true *p* value is .00001, but it would round to *p* = .000), you need to write it down as *p* < .001. The *p* needs to be printed in italic always.

Chapter 6: Results

- Consider showing model fit

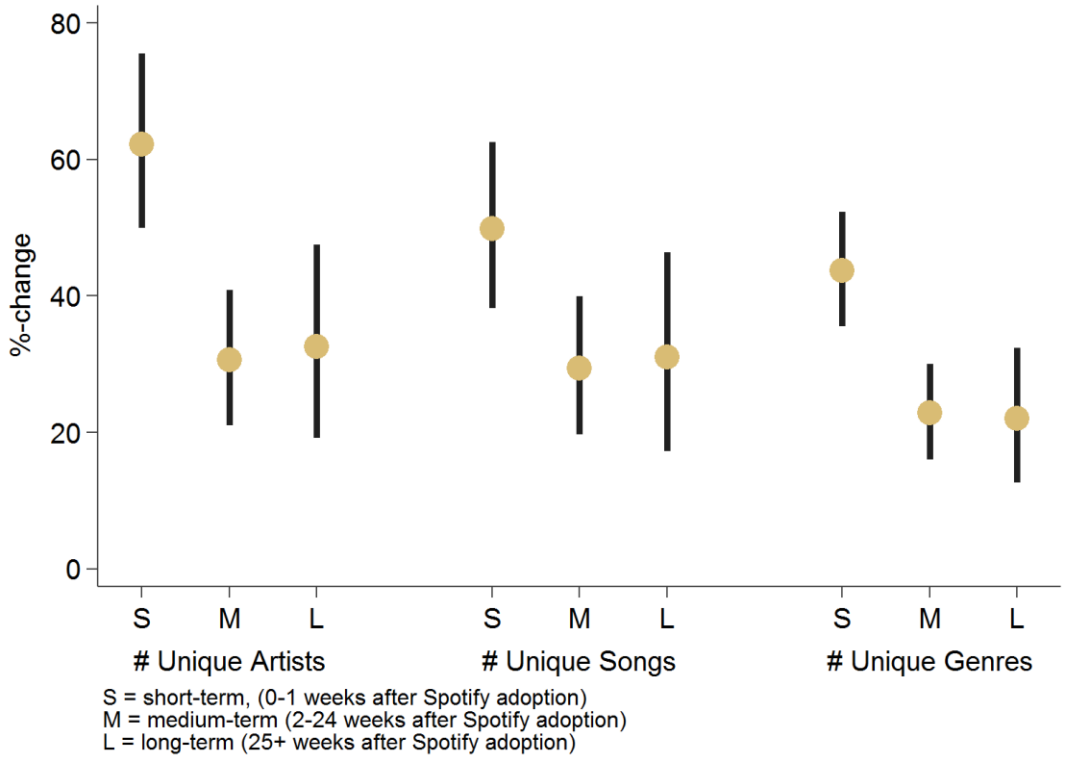
Figure 2
MODEL FIT



Datta, H., Foubert, B., & Van Heerde, H. J. (2015). The challenge of retaining customers acquired with free trials. Journal of Marketing Research, 52(2), 217-234.

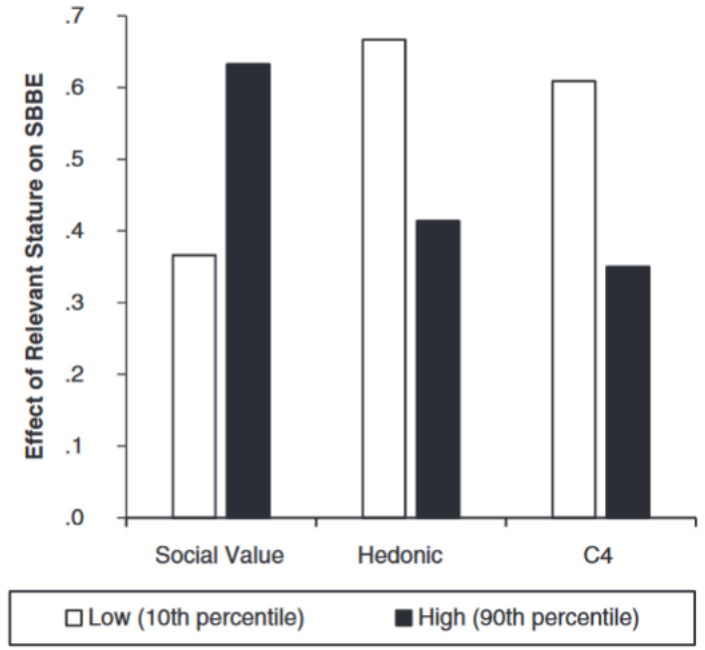
Chapter 6: Results

- Consider plotting some results. E.g., coefficients of your estimated models, including error bars.



Datta, H., Knox, G., & Bronnenberg, B. J. (2017). Changing their tune: How consumers' adoption of online streaming affects music consumption and discovery. Marketing Science.

FIGURE 3
Effect of Relevant Stature and Energized Differentiation on SBBE for Different Levels of Moderators

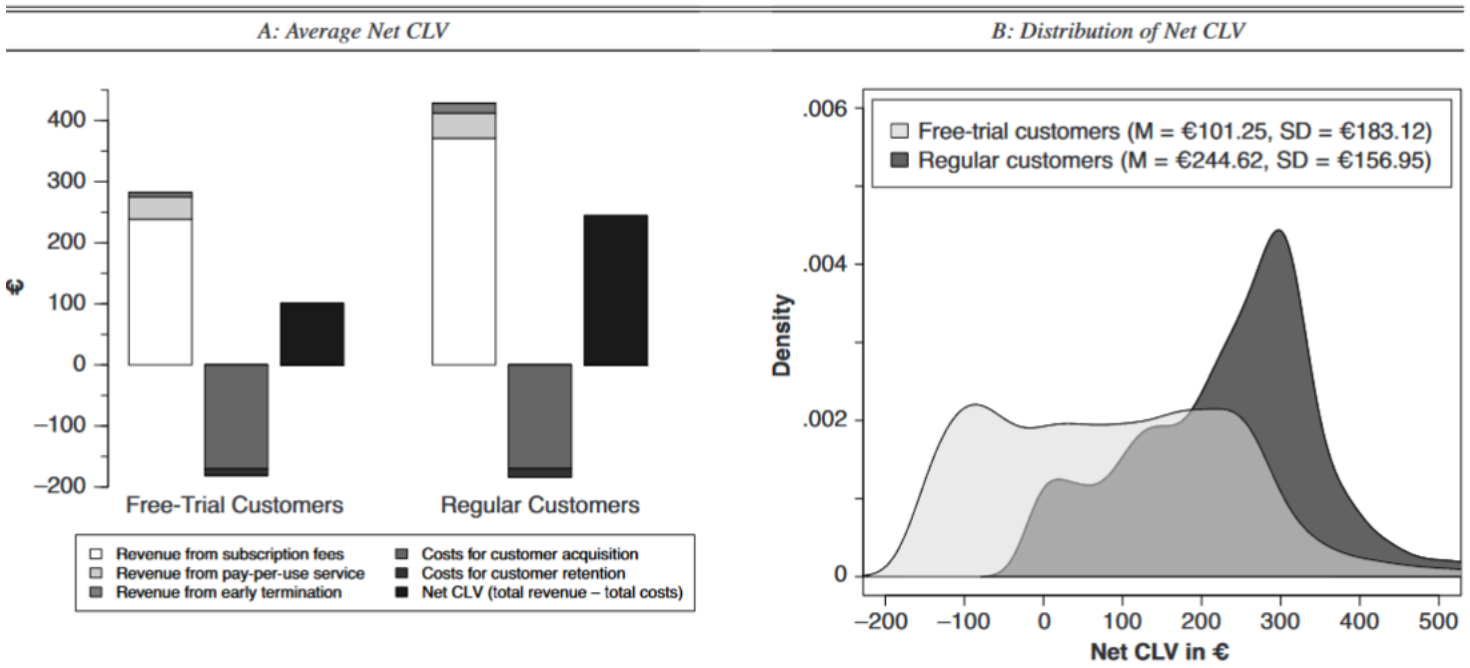


Hannes Datta, Kusum L. Ailawadi, and Harald J. van Heerde (2017) How Well Does Consumer-Based Brand Equity Align with Sales-Based Brand Equity and Marketing-Mix Response?. Journal of Marketing: May 2017, Vol. 81, No. 3, pp. 1-20.

Chapter 6: Results ("What-if" scenarios)

- Having estimated a regression model, you can come up with "what-if" scenarios. Consider plotting the outcomes or showing them in tables.
- As an example, see the outcome of a simulation of what free trial versus regular acquisition does to a customer's CLV.

Figure 4
NET CLV FOR FREE-TRIAL AND REGULAR CUSTOMERS



Datta, H., Foubert, B., & Van Heerde, H. J. (2015). The challenge of retaining customers acquired with free trials. Journal of Marketing Research, 52(2), 217-234.

Chapter 7: Discussion

- Summary of main findings
  - Summarize in a concise way
  - You can also have a summary table with your results (e.g., hypotheses/expectations confirmed/not confirmed, etc.)
  - Some inspiration → see table on the right
- Theoretical and managerial takeaways
  - What can researchers do with your findings? What can managers do and how do you think will it influence their actions?
- Limitations and future research
  - No research is perfect. Neither is yours. Be open about its shortcomings!

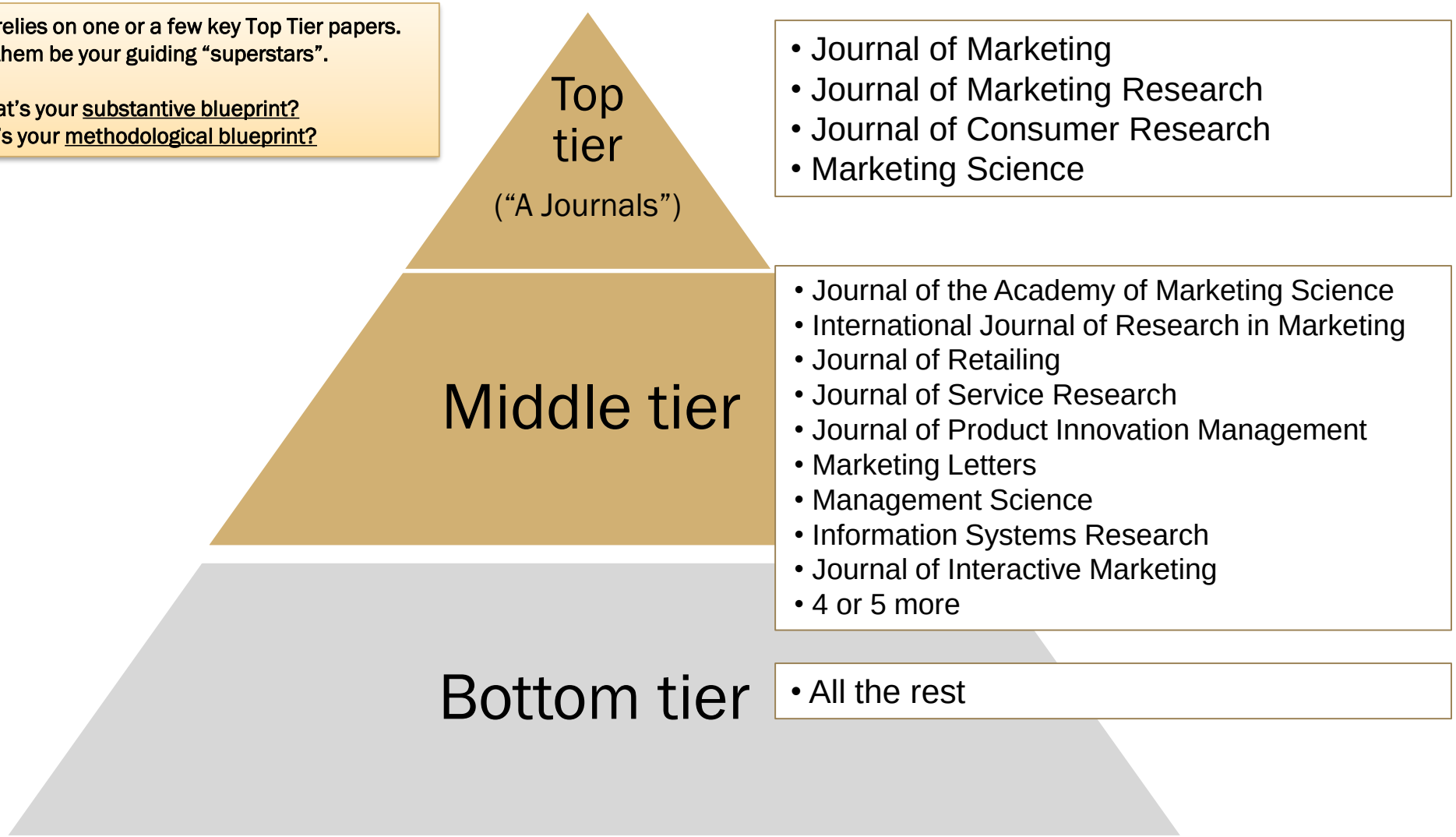
TABLE 7
Association of CBBE with SBBE and Marketing-Mix Elasticities: Summary of the Findings

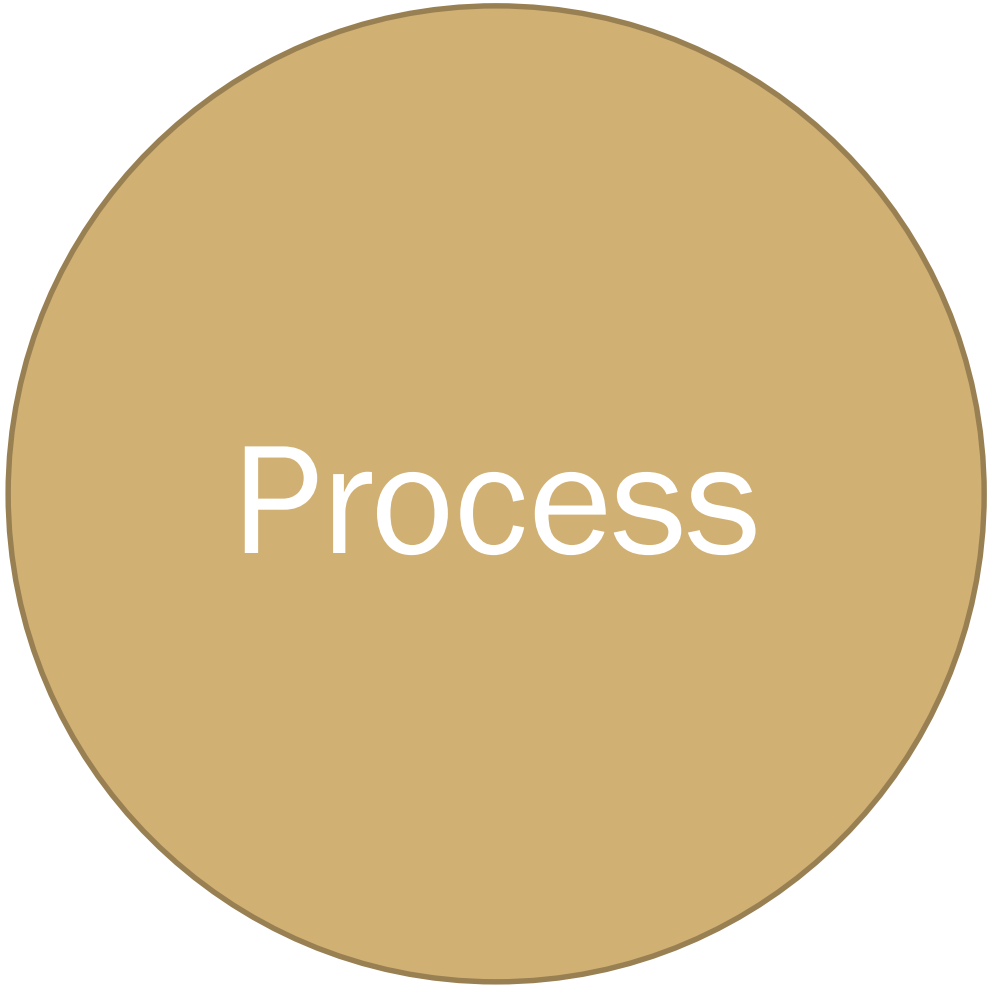
CBBE Dimension	Association With... SBBE	Association With... Marketing-Mix Elasticities
Relevance, esteem, knowledge, combined in relevant stature	• There is positive and significant correlation between SBBE and relevance (.39), esteem (.35), and knowledge (.53). • The effect of relevant stature on SBBE is significantly positive (P_1 supported). • The effect of relevant stature on SBBE is stronger for ◦ Less concentrated categories ($P_{1.2}$ supported). ◦ High–social value categories ($P_{1.3}$ supported). ◦ Less hedonic categories ($P_{1.4}$ not supported). • The effect of relevant stature on SBBE is not significantly moderated by functional risk ($P_{1.1}$ not supported).	• Higher relevant stature is associated with ◦ No significant difference in regular price elasticity (P_3 not supported). ◦ A stronger (more negative) promotional price elasticity (P_5 supported). ◦ A stronger feature/display elasticity (P_7 supported). ◦ A weaker distribution elasticity (P_9 no prediction). ◦ A stronger advertising elasticity (P_{11} supported).
Energized differentiation	• There is negative and significant correlation between SBBE and energized differentiation (–.14). • The effect of its principal component on SBBE is significantly negative (P_2 supported). • The effect of energized differentiation on SBBE is stronger for ◦ More concentrated categories ($P_{2.2}$ supported). ◦ More hedonic categories ($P_{2.4}$ supported). • The effect of energized differentation on SBBE is not significantly moderated by ◦ Functional risk ($P_{2.1}$ not supported). ◦ Social value ($P_{2.3}$ not supported).	• Higher energized differentiation is associated with ◦ No significant difference in regular price elasticity (P_4 not supported). ◦ A weaker (less negative) promotional price elasticity (P_6 supported). ◦ No significant difference in feature/display elasticity (P_8 not supported). ◦ No significant difference in distribution elasticity (P_{10} not supported). ◦ A stronger advertising elasticity (P_{12} supported).

Tip: Which journals to use?

A good thesis relies on one or a few key Top Tier papers.
Make them be your guiding "superstars".

What's your substantive blueprint?
What's your methodological blueprint?




Process

My expectations

- Be active and prepared
- If you are supervised in a team, please send your material to each other early enough, so that you can (a) study each other's work, and provide (b) a round of feedback before submitting material to your supervisor
- Make appointments in advance
- If you experience any kind of problem, let me know so that we can find a solution
- Individual meetings are possible (but, typically not at the beginning of your thesis trajectory, max. 3)

Getting feedback

- I will actively share positive and negative comments about your work.
- The negative comments typically fall in one of the three categories below – which helps you to gauge how important it is to address them in a revision

Critical	• You really have to address the comment! • E.g., you did a mistake.
Disputable	• You disagree with the comment, but you still have to convince your advisor/reviewer about it. • E.g., you insufficiently explained your idea
Gentle	• Thank your supervisor/reviewer for the comments… and that's all about it. The comment is not worth to be addressed, but be polite about it. • E.g., …hm… I won't say now :)

Managing me (i.e., your advisor)

- Be critical and take responsibility
  - This is your project, not mine! Even if I gave you some data, and even if you are working on topics related to my own research interest…
- Prepare thoroughly for meetings
  - Have an agenda of things you need to discuss
  - Send stuff early enough to allow me to prepare
  - Not only come up with problems, but also with solutions
- During our meetings
  - Don't take criticism personally (I'm just trying to improve your work)
  - If you don't know the answer to a question on the spot, say that you need time to think about it
- Keep me up-to-date
  - Provide a small recap of your project regularly

Managing me (i.e., your advisor)

The don'ts

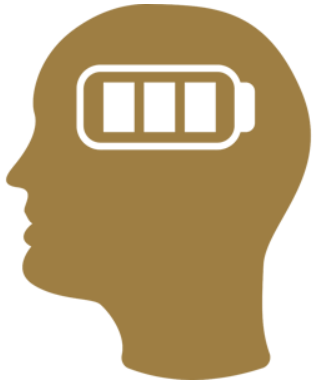
- Don't stay away for many weeks without informing me
- Don't contact me for every tiny question you have, but first try to find a solution yourself or in discussions with your thesis partner
- Don't plagiarize

The do's

- Do make an appointment before you come see me
- Do send me a reminder if I do not reply to your email within three working days
- Do submit things on time and stick to agreements
- Do take into account my agenda
- Do accept my feedback, but also give your feedback to me (!)
- Do write code/syntax (e.g., in SPSS, R)
- Write short and to-the-point emails
- If you call me (e.g., Skype/Hangouts), be prepared to share your screen. I prefer also talking with a webcam. And btw, please test your audio, too!

Keeping your sanity

- Follow your own pace: don't compare to others
- Don't think you haven't done anything
- Realize that ups and downs are inevitable
  - E.g., questionnaires, data preparation and analysis
- Do things not related to your thesis
- It's good to be ambitious (I am, too!), but realize that your thesis most likely will not change the world
- Talk to other students of mine / form groups



writing and form

Write actively, and engage your readership

- Write in active voice (as if you're having a conversation with a good academic friend)
  - Example 1
    - Wrong: "The data has been analyzed and the following results found"
    - Correct: "We analyze the data, and find that [...]"
  - Example 2
    - Wrong: "three constructs have been combined"
    - Correct: "we combine three constructs"

- Use the collective "we" (or "I")
  - Don't be shy! Either use "I" or the collective "we" – this will help you to tell a better story, and also avoid writing in passive voice.

- Avoid writing in "constructs", but give meaning to everything you say
  - Wrong: "The overall click-through rate for users in the "promotion" condition is [...]"
  - Correct: "Promotion-focused users had higher click-through rates than [...]"

Say things efficiently

- Be brief
  - Perfection in writing
    - …can be defined as ***"There's nothing else left in my thesis to take out".***
    - …and certainly is not ***"There's nothing more I can add to my thesis".***
  - Please replace wordy constructions by shorter words ("say it shorter"); see also writing guide
- Avoid direct citations, unless there is really no better way to say it
  - Wrong: Thomson (2009) describes these labels as: "A label that is independently owned and seems to be responsible for its own radio promotion"
  - Correct: According to Thomson (2009), these labels are independently owned, and responsible for promoting their content to radio stations themselves.

Avoid typos and grammar mistakes

Some students seemingly have established a dangerous mindset about the details…

- "Missing, incomplete or wrong references are okay in a master thesis."
- "Punctuation is no big deal."
- "Grammar issues can be addressed later."
- "Details in figures and tables are only relevant in the final version of the thesis."
- "The listed maximum length of a thesis is really only a rough guideline; it is okay if the manuscript is actually longer."

Recommendations:

- Pay attention to the small details; do not allow issues with the details to get in the way of an otherwise good thesis.
- Ask friends and fellow students for feedback & proof-reading! Do so many times!
- Refer to the '"Academic's Little Helper" (tiu.nu/writing) and my own papers, available on my website (datta-online.com) and my Tilburg profile (tiu.nu/datta).
- If writing remains an issue for you, please be in touch with the Scriptorium – they can help you greatly!

Imitate a style

- Imitate my style of writing (or the style of any of your favourite authors in your field of research).

- A “paper format” (like you find in the journals) overrules the ‘traditional’ thesis format as communicated in the Master thesis guide by our department.

Formatting your document

- Format of your thesis
  - A4, 1.5-spaced
  - Times New Roman, 12pt.
  - Tables in 8pt or 10pt. Footnotes in 10pt. No exception. No fancy title pages.

- Use page numbers (bottom right)
  - same font as your text, 10pt
  - page numbers start with 1 on the first page after your title page (the title page doesn't have any page number)

- Adhere to the English way of formatting numbers
  - Decimal *points*, not commas (e.g., the average is 10.52, and not 10,52)
  - Separation of groups of digits (e.g., there are 1,531 users in the data set, and not 1531; similarly, there are 1,531,213 song plays, instead of 1531213 song plays)
  - Drop leading zeros (e.g., .15 instead of 0.15)

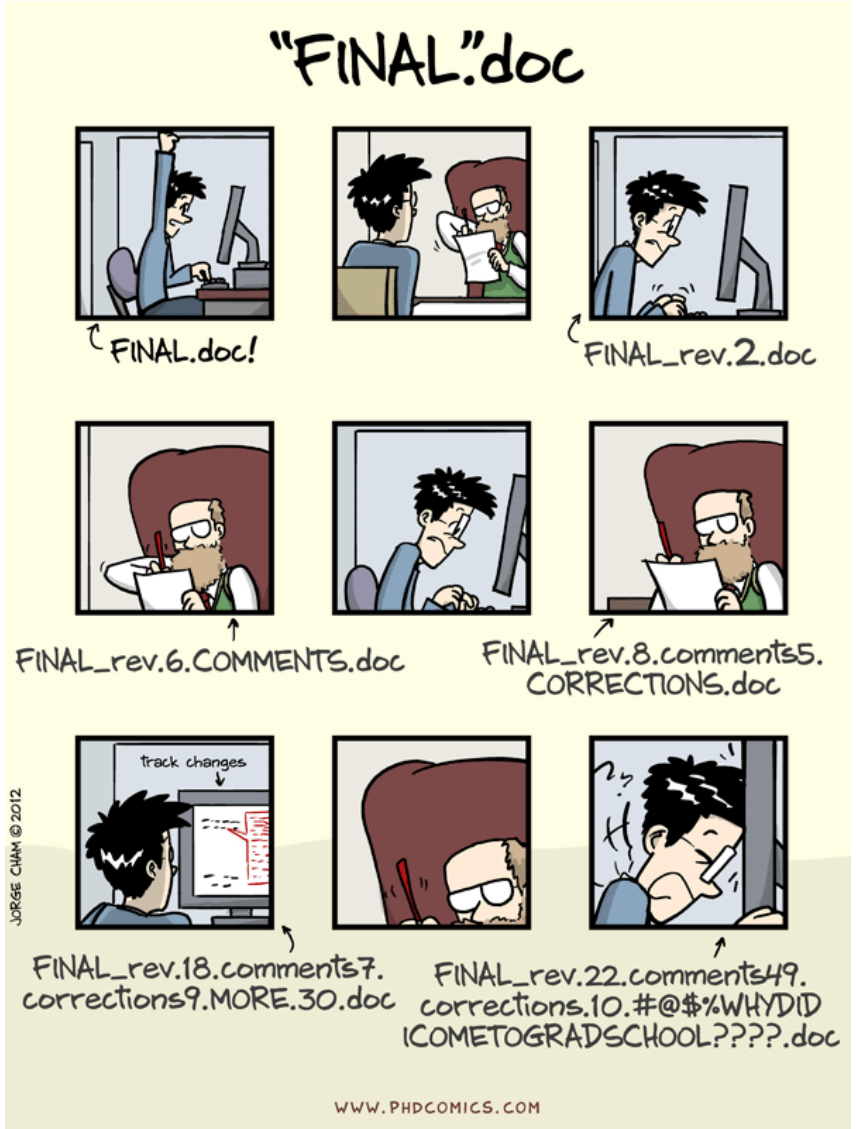
- Take time to proof-read your thesis

Table and figure formatting

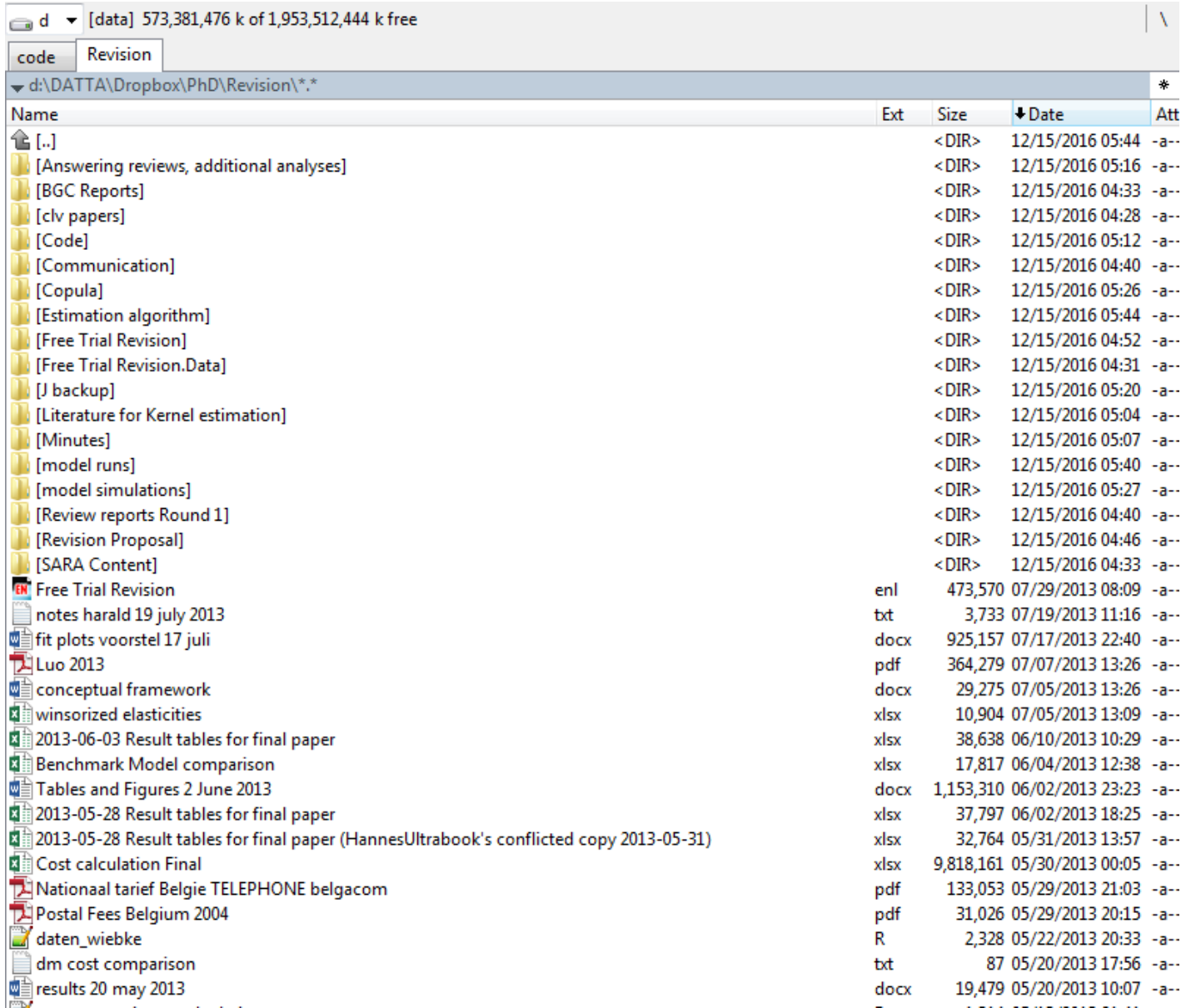
- Tables need to be formatted yourself, e.g., in Word or Excel,
- You are not allowed to just paste the output from your statistical program in the main text (in the Appendix, it's ok!)
- Tables (and also figures) should be understandable without reading the corresponding sections (i.e., use clear titles, and add notes wherever necessary)
- Note that you can also combine tables into larger ones, and thereby save space (e.g., tables with the same explanatory variables but different coefficients)
- Avoid colors (also in figures!) – I'll print your thesis in greyscale most of the time.

stay organized

Final.docx...


"FINAL".doc
FINAL.doc!
FINAL_rev.2.doc
FINAL_rev.6.COMMENTS.doc
FINAL_rev.8.comments5.
CORRECTIONS.doc
track changes
FINAL_rev.18.comments7.
corrections9.MORE.30.doc
FINAL_rev.22.comments49.
corrections.10.#@$%WHYDID
ICOMETOGRADSCHOOL????.doc
JORGE CHAM © 2012
WWW.PHDCOMICS.COM

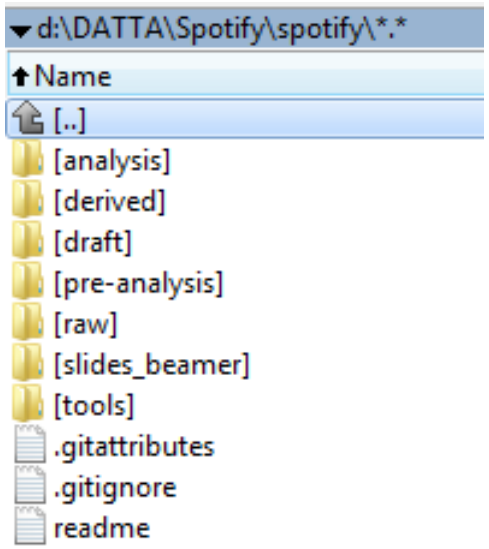
My first own project in my PhD is a messed up directory...!



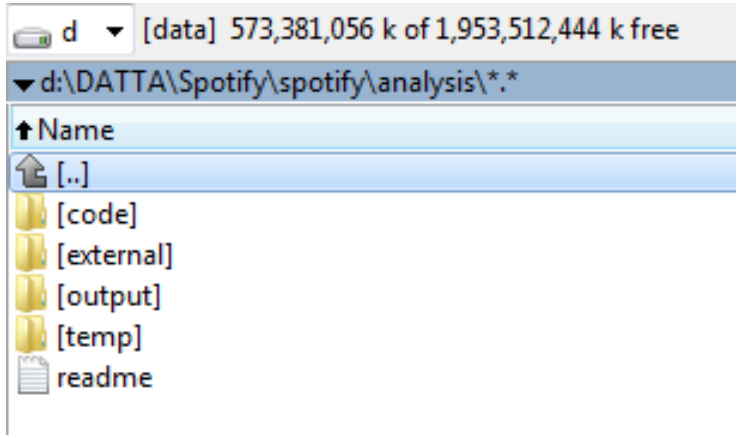
Name	Ext	Size	Date	Att
[..]		<DIR>	12/15/2016 05:44	-a--
[Answering reviews, additional analyses]		<DIR>	12/15/2016 05:16	-a--
[BGC Reports]		<DIR>	12/15/2016 04:33	-a--
[clv papers]		<DIR>	12/15/2016 04:28	-a--
[Code]		<DIR>	12/15/2016 05:12	-a--
[Communication]		<DIR>	12/15/2016 04:40	-a--
[Copula]		<DIR>	12/15/2016 05:26	-a--
[Estimation algorithm]		<DIR>	12/15/2016 05:44	-a--
[Free Trial Revision]		<DIR>	12/15/2016 04:52	-a--
[Free Trial Revision.Data]		<DIR>	12/15/2016 04:31	-a--
[J backup]		<DIR>	12/15/2016 05:20	-a--
[Literature for Kernel estimation]		<DIR>	12/15/2016 05:04	-a--
[Minutes]		<DIR>	12/15/2016 05:07	-a--
[model runs]		<DIR>	12/15/2016 05:40	-a--
[model simulations]		<DIR>	12/15/2016 05:27	-a--
[Review reports Round 1]		<DIR>	12/15/2016 04:40	-a--
[Revision Proposal]		<DIR>	12/15/2016 04:46	-a--
[SARA Content]		<DIR>	12/15/2016 04:33	-a--
Free Trial Revision	enl	473,570	07/29/2013 08:09	-a--
notes harald 19 july 2013	txt	3,733	07/19/2013 11:16	-a--
fit plots voorstel 17 juli	docx	925,157	07/17/2013 22:40	-a--
Luo 2013	pdf	364,279	07/07/2013 13:26	-a--
conceptual framework	docx	29,275	07/05/2013 13:26	-a--
winsorized elasticities	xlsx	10,904	07/05/2013 13:09	-a--
2013-06-03 Result tables for final paper	xlsx	38,638	06/10/2013 10:29	-a--
Benchmark Model comparison	xlsx	17,817	06/04/2013 12:38	-a--
Tables and Figures 2 June 2013	docx	1,153,310	06/02/2013 23:23	-a--
2013-05-28 Result tables for final paper	xlsx	37,797	06/02/2013 18:25	-a--
2013-05-28 Result tables for final paper (HannesUltrabook's conflicted copy 2013-05-31)	xlsx	32,764	05/31/2013 13:57	-a--
Cost calculation Final	xlsx	9,818,161	05/30/2013 00:05	-a--
Nationaal tarief Belgie TELEPHONE belgacom	pdf	133,053	05/29/2013 21:03	-a--
Postal Fees Belgium 2004	pdf	31,026	05/29/2013 20:15	-a--
daten_wiebke	R	2,328	05/22/2013 20:33	-a--
dm cost comparison	txt	87	05/20/2013 17:56	-a--
results 20 may 2013	docx	19,479	05/20/2013 10:07	-a--

That's why… this is the directory structure you should use

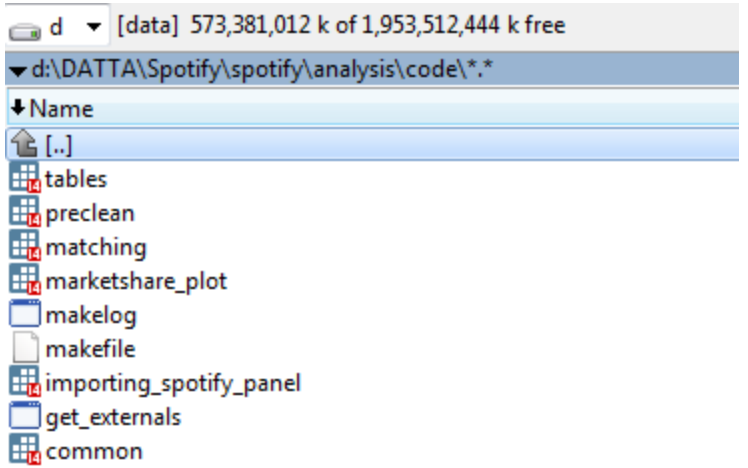
Main directory of a project



Taking a closer look at \analysis



Taking a closer look at \analysis\code



Keeping things clean and readable

- Understand things months/years after you did them
- Let others understand your code (e.g., comply with *Marketing Science*'s replication policy; get cited more frequently)

Option 1: Use version control

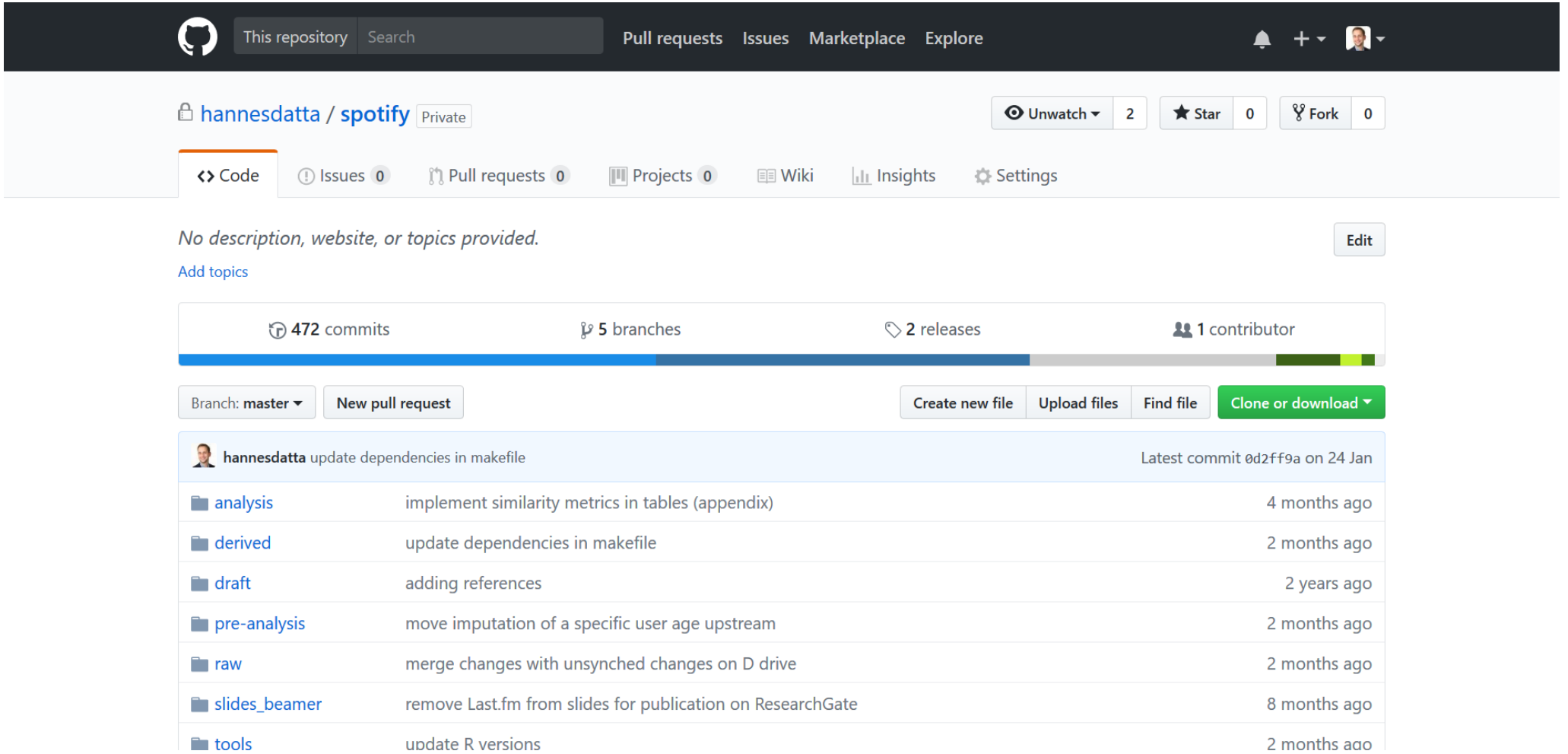
- (1) Unlimited “undo”
  - Which means you will delete unnecessary stuff (but you can recover it any time)

- (2) Collaboration
  - Work in parallel (vs. “George has unsaved changes”, better than emailing)
  - Resolve conflicts (vs. “Hannes’s conflicted copy”)
  - Applies even to collaboration with yourself (at home, on campus, etc.)

→Gain productivity (and I really mean it!!!)

More info? See github.com and https://software-carpentry.org/lessons/

Option 1: Use version control (GitHub project management)



Option 2: Use cloud storage (e.g., Google Drive)

- Make regular backups!
  - TilburgU has a free, unlimited Google Drive account for you…
  - Dropbox also works its magic…!

- Even if not using version control, always adhere to the directory structure given to you.

Tips for our meetings

First meeting with supervisor

- Send research proposal to your thesis partner.
  - If you have updated your proposal after you uploaded it, please also send a copy to me at least three days before the meeting.

- Make an overview of the data that you have or plan to use
  - e.g., inform yourself about the dataset, make some descriptive statistics if possible; see “describe the raw data” in this booklet/slide deck

- Make a time planning (including deadlines), and an envisioned date of your defense.
  - This is typically your first meeting + 4 or 5 months. Check the department guidelines on this. We will fix your latest defense data when we meet.

- We will most likely discuss the contribution of your study, and your data.
  - Study these sections well and be prepared to pitch your research!!!
  - Have a list of potential contributions of your research
  - Have an initial conceptual model

If these things are not part of your research proposal (yet), prepare them, please!

Next meetings with supervisor

- Follow the guidelines from this booklet/presentation
- Plan enough time to collect, and get a feel for the data!
  - This also holds for data clean-up, and analysis
- Read my papers to understand the “style” of research I do
- Probably we will be iterating through this phase for some time before we start with data analysis, etc.

Your defense

- Plan your defense about 4 weeks in advance; your thesis advisor will contact your co-reader to coordinate about the date
- Hand in a copy of your thesis to the mailboxes of your thesis advisor and co-reader a week before your scheduled defense date, and upload your thesis to MaMa
- During your defense, you will be asked to clarify your thesis by answering questions. You will not have to present your work to us (e.g., by means of a PowerPoint presentation)
- Please bring your own copy of the thesis; we may refer to it during our meeting and it is useful to take a look at the tables/figures/paragraphs

Contact

Hannes Datta
T606 (Tias Building)

Email: h.datta@tilburguniversity.edu
Book Appointments via: calendly.com/hannesdatta

Social:
https://www.linkedin.com/in/hannes-datta
@hannesdatta
Whatsapp for Business: +31 13 466 8938 (you can add me using your regular Whatsapp account)