QBUS6860 Visual Data Analytics

Final Assignment Report Example

This is an example the final assignment report for the QBUS6860 unit (Visual Data Analytics). **This example is intended to give you a sense of the scope of the final assignment. Note that this analysis is far from perfect.** Comments give you an idea of things, which did not work well in this report. As you read through, ask yourself what you might do differently! Do not take this example as something you would want to replicate as **it will not guarantee you an excellent mark**. Please, also keep in mind that apart from the report, you need to submit the following:

1. Video recorded presentation (3-5 min) summarising your analysis, findings,

   a. As a file to Canvas in one of these 4 formats: MP4, MOV, WMV and AVI
   b. As an Unlisted (non-searchable) and unlisted (unsearchable) link to your video on one of these platforms: YouTube, Vimeo, Bilibili

2. Written report of not more than 5,000 words (excluding references) - this should be your presentation script together will all graphs and any additional materials or outputs of your analysis, which you would like to include.

3. Your Python code (preferably in .html format)

4. If you make significant changes to data provided for the final assignment (e.g., if you mined additional data, if you merged several .csv files provided, etc.) you also must provide your dataset so that we could check correctness of your analysis. If you use one of the .csv files provided "as is", you will not need to submit your data.

QBUS6860 Visual Data Analytics

Final Individual Assignment

Title: **The Impact of Big Data Analytics on Innovation in Formula 1**

Student Number: []

Assigned visualisations: **bar chart** and **line chart**

Word Count: 3714 (excluding References, Tables and Figures)

Commented [GP1]: Do not forget to include your student number

The Impact of Big Data Analytics on Innovation in Formula 1

1. Background

The challenges and opportunities posed by Big Data have attracted the attention of both scholars and industry practitioners (Akhtar *et al.*, 2017; Marr, 2015; Reinmoeller and Ansari, 2016). Academic research acknowledges the opportunities offered by Big Data when information (data) is translated into decision making strategies and improved innovation and performance (Chen, Chiang, and Storey, 2012; McAfee and Brynjolfsson, 2012; McKinsey Global Institute, 2011). Current debate also points to the competences companies need to deal with advanced technology and Big Data (Akhtar *et al.*, 2017; Reinmoeller and Ansari, 2016). Although the benefits and challenges identified in the management literature are numerous, the link between Big Data and innovation remains largely anecdotal due to lack of empirical work on how large datasets can influence business outcomes (Kache and Seuring, 2017; Sen, Ozturk, and Vayvay, 2016). Several scholars have proposed frameworks for how Big Data applications can be exploited to generate value, relying mainly on a case-based research methodology (Matthias *et al.*, 2017). However, the generalizability of these findings to a wider population of companies is difficult. The lessons learnt may be unique to the *in-situ* performance at a particular time. A systematic review conducted by Frizzo-Barker et al. (2016), of Big Data papers published between 2009 and 2014, acknowledges the lack of empirical work and shows that this stream of research is dominated by conceptual papers.

Commented [GP2]: This background is all right, but it would be nice to add how important is F1 industry, what is the size of it? Why it is important to look at F1 industry?

2. Business Question and Question Justification

This report addresses the following business question: ***what is the impact of Big Data analytics on innovation in Formula 1*?** This question is not only applicable to the Formula 1 industry, but also to other analytically dense industries. Yet, F1 constitutes an ideal setting to investigate the use of Big Data (Aversa, Cbantous, and Haefliger, 2016; George, Haas, and Pentland, 2014). First, in the F1 context, major innovations are distinctly and precisely measured (Gino and Pisano, 2011). Furthermore, unlike many industries (where products/services are heterogeneous), all F1 teams produce homogeneous output (final standing in the race ranking), which allows us to compare team performance directly and more precisely (Goodall and Pogrebna, 2015). Second, F1 is a highly analytically dependent and data-dense industry and has seen a transition from systematic manual processing of data to predictive analytics and, more recently, evolution to a mature stage in the Big Data revolution. Their performance depends on how the teams respond to these data, which makes a good testbed for information processing theory (Rogers, Miller, and Judge, 1999; Tushman and Nadler, 1978) operationalized in the Big Data operations management domain. Third, in most high technology industries, the strictness of safety and regulatory standards increases over time as more information is revealed. Examples are the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) and Restriction of Hazardous Substances (RoHS) regulation, which is aimed at tighter control of EU supply chains by monitoring substances used in products (Westervelt, 2012). This tighter regulatory control involves increasingly more complex data collected from manufacturers. The F1 industry has experienced the imposition of many regulations over time (Marino *et al.*, 2015) and how teams respond and adapt to process 'future data' into existing systems before making an informed decision about the design of cars and race strategies provide important lessons for other industries.

Commented [GP3]: This is a good question without an obvious answer. This section provides good justification of the question and why F1 is an appropriate context to study it.

Scholars have recognized the distinctive characteristics of F1, for example, Jenkins (2010) suggests that F1 firms possess sustaining capabilities - munificent resource configurations that extend the time available for firms to adapt to technological changes – thereby allowing them to remain competitive across discontinuities. In this study, I focus on the Big Data structure (BDS) of F1 teams from the perspectives of information integration and information accessibility. Building on an information processing perspective, which emphasizes the need to process information by considering external demand (Rogers *et al.*, 1999; Tushman and Nadler, 1978) and the technological environment in which firms operate (Hughes *et al.*, 2014; Hughes, Hughes, and Morgan, 2007), I examine how different strategies for Big Data information processing lead to different BDS configurations. I show how different BDS structures influence team performance in terms of output (achievement of podium positions) as well as innovation production.

3. Methodology, Hypothesis and Justification of Selected Analytical Tools

F1 is recognized as a unique setting to investigate the role of Big Data and business analytics, since it relies heavily on sophisticated applications of real-time information systems to support informed decision making processes during a race (Aversa *et al.*, 2016; Aversa, Furnari, and Haefliger, 2015; George *et al.*, 2014; Goodall and Pogrebna, 2015; Marino *et al.*, 2015). F1 is estimated to be worth approximately $6 billion annually (Sylt and Reid, 2011). Constructor teams' profits come from advertising and TV. A higher finishing position, primarily a podium position (1st to 3rd), generates more sponsorship and TV income. Increasingly, modern teams are raising money from the development of F1 technologies that spill over into other industries. For example, Williams and McLaren (Applied Technologies) have associate companies. It is an interesting industry intellectually because it is subject to a great deal of regulatory turbulence. The Fédération Internationale de l'Automobile (FIA), the F1 industry governing body, imposes strict conditions, which are revised annually, on all aspects of F1 (the teams, technology, resources, track, tyres, drivers, etc.). The link between regulation and innovation has been well documented (Stewart, 2010). *T is* embodied in F1; regulation is unambiguously associated with innovation and performance (Jenkins, 2004, 2010; Jenkins, Pasternak, and West, 2007; Khanna, Kartik, and Lane, 2003; Marino *et al.*, 2015) and regulatory compliance results in a level playing field for all competing teams. Indeed, sometimes rule changes are made with the specific intention of curtailing the dominance of one team, for example, Ferrari and Michael Schumacher in 2003 (Hoisl, Gruber, and Conti, 2017).

Commented [GP4]: I guess you mean "It is" – be sure to check your work for typos before you make your submission

F1 is an extremely data-dense industry with sophisticated data analytics. All contemporary F1 cars are using sophisticated *telemetry* systems to obtain, transmit, process, and analyse information. According to NASA, the term *telemetry* originates from the Greek "tele" which means "remote" and "metron" which refers to "measure" and depicts a process of automatized communications and transmissions allowing to obtain data from remote or poorly accessible points for monitoring (SAO/NASA Astrophysics Data System, 1987). In the F1 context, telemetry is implemented through a large number of sensors and electronic devices inclusive of the Electronic Control Unit (ECU) which communicates data to the pit wall, pit garage or another remote site (e.g., Toet, 2013).

Each Formula 1 car is equipped with 150-300 sensors (dependent on the racetrack, weather conditions, and other factors). For the reasons explained below, F1 telemetry is a one-way transmission system: the data is sent from the car to the engineering and strategy team, but the team does not have an opportunity to send the data to the car. According to various technical forums[1], the data from the car is transmitted wirelessly using 1,000-2,000 encrypted telemetry channels using either 1.5 GHz frequency or another locally allowed frequency. While the delay between the data collected and received at the team boxes varies, on average, it is around 2 milliseconds. Since the received data is compressed, the number of actual gigabits of

[1] See, e.g., the technical forum at http://www.formula1-dictionary.net/telemetry.html for more detail.

data received by teams may differ from race to race, although each team collects approximately 1.5 billion samples of data from a single race and approximately 5 billion of samples throughout the race weekend (this includes data from all training sessions). The transmitted data on engine performance, suspension state, gearbox performance, fuel status, temperature readings including tires temperature, g-forces and actuation of controls by the driver is analysed by the engineering and strategy team and results of this complex live analytics is communicated to the driver in the form of racing strategy advice.

Since telemetry is the major source of (live) Big Data for the F1 teams, in my analysis I use the telemetry development stages in the F1 industry as a "natural" proxy of (Big) data analytics evolution. Since 1950s, an F1 team performance was highly dependent on this team's ability to collect, process, and analyse large amounts of data. Historically, we can distinguish between 6 phases in the F1 data analytics progress using the development of telemetry as a proxy for determining the boundaries between different stages. The Big Data analytics history of the F1 industry proxied through the evolution of the use of telemetry was drawn from Jenkins (2010), the data collected by the McLaren F1 team[2] as well as from the Formula 1 Dictionary Technical Forum.[3]

I distinguish between the following phases in F1 data analytics (depicted in Table 1). The evolutionary phase-based approach summarised in Table 1 allows us: (i) to identify the pre-telemetry period (Phase 1); (ii) to identify phases of significant heterogeneity between F1 teams in terms of their access to telemetric technology (Phase 2 and Phase 4); as well as (iii) to understand when F1 teams had similar or standardized telemetric technology (Phases 3, 5, and 6). Using the identified phases, I can now use F1 performance data to understand the differences between phases and explore whether performance of industry as a whole as well as performance of individual players changed from one phase to the other.

To develop my hypothesis about innovation in Formula 1, I use a combination of Linden and Fenn (2003) and Fenn and LeHong (2011) Gartner Hype Cycle framework. I assume that analytics phases shown in Table 1 determine the boundaries of Gartner Hype Cycle stages of innovation development: Technology Trigger, Peak of Inflated Expectations; Through of Disillusionment; Slope of Enlightenment; and Plateau. Therefore, my main **hypothesis** is that ***the lifecycle of major innovations in Formula 1 follow the Gartner Hype Cycle shape where each stage is determined by the data analytics phase from Table 1.*** Specifically, I hypothesise that the correspondence between analytics and innovation will follow the pattern depicted on Figure 1. Specifically, Phase 1 (Driver as a Sensor) pre-dates telemetry analytics and, therefore, represents a preliminary stage. This stage leads to Phase 2 (Telemetry Development) where the development of telemetry technology is triggered which causes as increase in the number of innovations. Phase 3 (Early Telemetry Phase) represents the peak of inflated expectations where data potential is uncovered by the industry and positive "hype" is created allowing to reach the peak of innovation. Phase 4 (Turbulence) captures a "through of disillusionment" where drawbacks of the telemetry technology start to significantly outweigh the benefits and number of innovations rapidly decrease. Phase 5 (Mature Telemetry) phase is equivalent to the slope of enlightenment where new capabilities of technology provide a new positive boost to innovation. This boost flattens or even disappears in Phase 6 (Big Data Telemetry) when technology reaches its plateau or even post plateau stage.

[2] See https://www.mclaren.com/formula1/team/a-brief-history-of-computing-in-F1-1052199/ for more detail.
[3] See http://www.formula1-dictionary.net/telemetry.html.

Table 1: Phases of Data Analytics in F1 Industry

Phase	Time period	Major milestones
1. Driver as a Sensor	1950-1974	The majority of teams used drivers as sensors who fed back the information about the car performance to the teams after the race.
2. Telemetry Development	1975-1988	1975 - McLaren started to experiment with telemetry first deployed 14 sensors on IndyCar. Until late 1980s – F1 teams started to experiment with telemetry.
3. Early Telemetry	1989-2001	By 1989 - F1 teams used “patched” telemetry transmitting data when cars came close to pits Early 1990s - F1 teams had high rate live information but it had blind spots (especially on tracks with dense trees or high buildings like Monza, Monaco, etc.). So information was incomplete. 1998 - Plextek[4] became a major supplier for telemetry systems 2000 - Incomplete information problem was fixed
4. Turbulence	2002-2004	2002 - Two-way telemetry was allowed (teams could not only receive but also send information to cars remotely) 2003 - FIA banned two-way telemetry
5. Mature Telemetry	2005-2012	2005 - Electronic Control Unit (ECU) TAG-310B SECU by McLaren Electronic Systems and Microsoft is developed 2008 - FIA standardized ECU for all F1 cars 2008 - FIA makes Advanced Telemetry Linked Analysis System (ATLAS) Express produced by McLaren Electronic Systems standard
6. Big Data Telemetry	2013-2020	2013 - Volumes of data become very high requiring major system upgrades 2013 - Upgraded standard ECU to the new TAG-320 SECU 2017 - Data Viewer developed by McLaren Applied Technologies[5]

[4] See https://www.plextek.com/.
[5] See https://www.mclaren.com/appliedtechnologies/

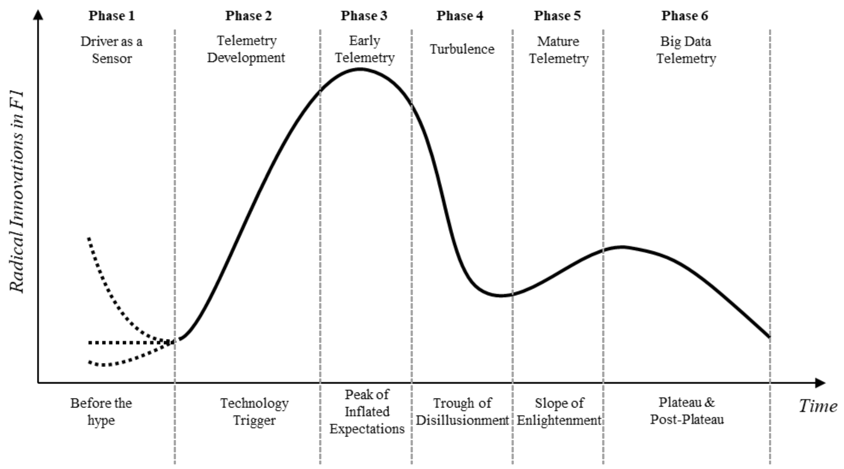


Figure 1: The Major Innovations Lifecycle in the Formula 1 Industry Predicted Based on the Phases of Data Analytics Evolution

Why wouldn't Big Data telemetry provide an extra boost to the innovation lifecycle in Formula 1 instead leading to a plateau? Information Processing Theory (IPT) provides a theoretical basis for this (Rogers *et al.*, 1999; Tushman and Nadler, 1978). Early research on IPT (Daft and Lengel, 1986; Daft and Weick, 1984) maintains that information gaps can be reduced by gathering more data. A large part of IPT discusses the reduction of uncertainty by facilitating decision makers' access to the right information at the appropriate time (Sakka, Barki, and Côté, 2016). However, obtaining more data is today not a major concern since most electronic devices and transactions generate abundant data. At the same time, greater access to data is linked to higher levels of misinformation and misinterpretation of those data. Although all F1 teams have access to Big Data, this does not mean that all of them will necessarily benefit from the data. In fact, greater volumes of information can be very difficult to process so only few market players may be capable of coping with the industry's increasing data supply.

In my analysis, I am planning to capture the major innovations' lifecycle using bar chart and provide several F1 teams use case illustrations using line chart. Instead of a single line chart, I am planning to use a chart with multiple lines, because it is more appropriate to demonstrate the differences between teams. Though bar chart is a good way to demonstrate life cycle, I will also use area chart in order to capture the dynamic shape better.

Commented [GP5]: This section could be expanded, and you can justify your visualisation choices better. Why are your visualisations appropriate? If you are using additional visualisations, what value do they add to your analysis? Please, explain.

4. The Data

Data used in this analysis spans 68 years from 1950 to 2017. The data includes 976 F1 races that took place in that period, with 22,083 car records forming our dataset. For each car record, I have data on: the starting and final positions of the cars that participated in each race, that is, team performance; constructor teams; their leaders' names, personal information and background; drivers' personal information and background; and information on each race circuit (weather conditions on the day, length of the circuit, etc.). The data were compiled from several sources (please, see my dataset attached as a separate file). For information on car

Commented [GP6]: Good, we need your data to check the correctness of your calculations/visualisations entries, circuit, constructor, drivers and other detailed Grand Prix race information, I used the provided dataset. Records in the starting dataset were cross-checked and augmented by other sources of data such as Wikipedia, Grand Prix Encyclopaedia, and other websites.[6] Data on F1 regulations were compiled from the website of the F1 regulatory body, Federation Nationale d'Automobile, https://www.fia.com/regulations.

Main dependent variable

My dependent variable is radical innovation. Radical innovation variable captures significant changes in technology which led to serious changes in the F1 industry (innovations associated with telemetry are excluded). In order to construct this variable, I used FIA regulations which were scanned for restrictions on or bans of major innovations complemented by data from a broad variety of technical online forums.[7] My dataset also includes the performance variable (final standings at the end of the race) for every race car in every Grand Prix season since the first year of the Formula 1 industry existence. I use the final race standings ranking as a determinant of performance for each car with 0 identifying the race winner; 1 – 2nd place; 2 – 3rd place; etc., to demonstrate the impact of Big Data on performance in a case study illustration.

Key Independent variable

My key independent variable is a nominal variable which identifies 6 phases of analytics development in the Formula 1 industry: Driver as a Sensor = 1; Telemetry Development = 2; Early Telemetry = 3; Turbulence = 4; Mature Telemetry = 5; and Big Data Telemetry =6.

5. Results

In the dataset, phases cover different number of years (see Table 2): specifically, while Phase 1 includes 25 years, Phase 4 has only 3 years. Yet, despite these differences the number of races per year increased in later years, meaning that there are at least 51 races in each phase providing sufficient amounts of data for our analysis.

I test my hypothesis by considering whether radical innovation in the F1 industry follows the Gartner Hype Cycle (see Figure 1). Since our data on radical innovation is annual data per team, we calculate the sum of all radical innovations per phase from all teams taking part in F1 Grand Prix competitions in that phase and then dividing the sum by the number of years in a given phase. Results of my calculations are presented in the 5th column of Table 2 as well as on Figure 2 (a) and (b). I show that innovations' lifecycle indeed follows the Gartner Hype cycle, where stages of the cycle are determined by the phases of data analytics development in the Formula 1 industry. This confirms my hypothesis.

Commented [GP7]: If you are using a provided dataset as a starting point or as your "complete" dataset, tell us what that dataset is. What csv file are you using exactly?

Commented [GP8]: Perhaps, at this stage you could have provided more detail on how you are planning to use performance variable exactly. So far, your research question and your hypothesis did not talk about performance.

Commented [GP9]: Should not these results be corrected for the number of seasons or total number of innovations (not just radical) per year? If not, you need to explain why.

[6] Data on radical innovations were collected from https://www.redbull.com; https://www.autosport.com;
http://damcdn.autosport.com; http://www.crash.net; www.racecar-engineering.com as well as other sources.
[7] Specifically, I used https://www.fia.com/regulations; https://www.redbull.com/us-en/5-technical-f1-innovations; https://www.jamesallenonf1.com/2017/02/five-of-the-best-f1-innovations-found-through-loopholes/; https://www.redbull.com/int-en/f1-best-designers-in-history; and https://www.autosport.com/f1/news/127947/ferrari-counting-on-piston-innovation.

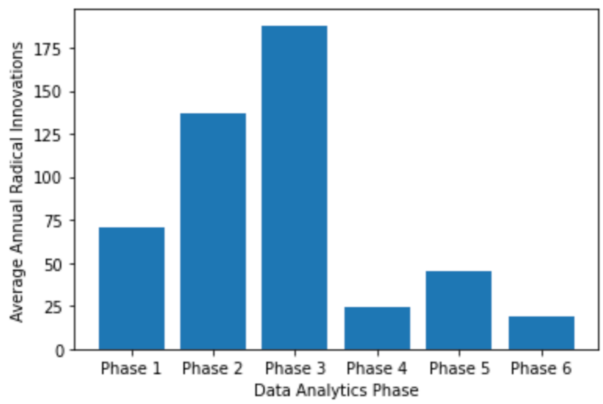
Table 2: Performance and Innovations by Phases of Analytics Development

Phase	Time Period	Number of years	Number of races	Average annual radical innovations	Normalised performance change
1. Driver as a Sensor	1950-1974	25	250	71	-0.10
2. Telemetry Development	1975-1988	14	218	137	0.03
3. Early Telemetry	1989-2001	13	212	188	-0.09
4. Turbulence	2002-2004	3	51	24	-0.18
5. Mature Telemetry	2005-2012	8	147	45	0.18
6. Big Data Telemetry	2013-2017	5	98	19	-0.13

Commented [GP10]: You have not explained how these results in this column were calculated.

My results suggest that analytical peak in terms on innovation in Formula 1 was during the Data Telemetry phase. And currently we are living through the data analytics plateau.

Commented [GP11]: This is way too short for results description. Make sure you explain what results you obtained and what they mean.



(a) Innovation and Data Analytics in F1 by Phase

(b) The Dynamics of Innovation in F1 by Phase

Figure 2: Gartner Hype Cycle for Radical Innovations in the F1 Industry

To illustrate my results, I will consider an example. I will consider the following well-known F1 teams: Ferrari, McLaren, Williams, Red Bull, Sauber, and Mercedes. On Figure 3 I show the historical normalised performance corrected for the number of major innovations of each of these 10 teams for all years where these teams took part in the F1 competitions. The value 0.5 on the vertical axis depicts average normalised performance so all values below 0.5 refer to poor performance relative to competitors while values above 0.5 capture successful performance relative to competitors.

Commented [GP12]: Why have you chosen these teams? You need to justify your choice.

Commented [GP13]: How did you calculate this? It is not clear, although I do like the example.

Figure 3 clearly shows team heterogeneity with regard to their ability to handle large amounts of data. The Ferrari team (one of two teams present in all 6 stages in the industry) exhibited a steady performance growth from Phase 1 to Phase 4 and became one of the top beneficiaries of the two-way telemetry[8]. It then suffered a slight decline in performance when the majority of teams gained access to telemetry in Phase 5 (the Mature Telemetry stage) and then managed to improve its performance through the use of Big Data telemetry in Phase 6. A completely different pattern is exhibited by McLaren who did not take full advantage of the two-way telemetry but then improved their performance in the mature telemetry stage. Yet, despite the strong analytics capability, McLaren seems to suffer from data overload in the Big Data phase where the performance of the team is significantly reduced. Interestingly, 4 teams saw their performance decline from Phase 5 to Phase 6: McLaren, Red Bull, Renault, and Sauber; one team did not significantly change its performance (Toro Rosso); while Mercedes, Ferrari, Force India, and Williams took advantage of Big Data. Haas team was only present in Phase 6 so it is not possible to make a comparison of performance between phases for this team. The biggest winner from the Big Data Telemetry appears to be Mercedes team whose performance received a significant boost from Phase 5 to Phase 6.

[8] See http://www.cds.caltech.edu/~murray/courses/cds101/fa04/caltech/17prix.html for more detail.

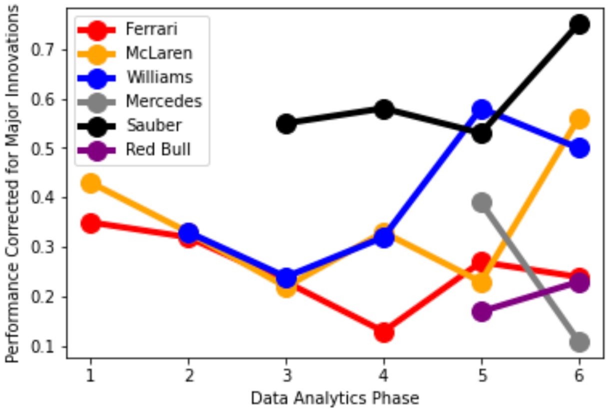


Figure 3 Historical Performance of F1 Teams By Analytics Phase Corrected for Major Innovations

Commented [GP14]: Note that normally you wouldn't want to this, but in this particular case, you do need to flip y axis. The lower is y, the better (as lower standing numbers 1,2,3 refer to better places in the race). The graph is a bit counter-intuitive.

Also, you need to explain the graph better. What does it show? What do we learn from it?

6. Insights for Business

This report set out to assess the impact of data analytics in general and Big Data analytics in particular on F1 teams' innovation as well as to provide empirical evidence to add to our understanding of current debate on the challenges and opportunities enabled by Big Data. Using real time telemetry evolution as a proxy of data analytics history in the F1 industry, I identified 6 phases of data analytics development. Using these phases, I showed that (1) data analytics phases as defined by telemetry technology evolution shape radical innovations in the F1 industry which follow the Gartner Hype Cycle; (2) performance in F1 follows a lagged Gartner Hype Cycle where performance suffers a delay compared to innovation; and (3) F1 teams exhibit significant heterogeneity in their ability to handle large datasets.

The findings from this study have important implications for managers working in an environment characterized by complex information and where the ability to process this information represents a distinctive competence which leads to better performance. I show that more data does not necessarily lead to more innovation or better performance. Rather, players who are capable of quickly adapting to changing data environment tend to win (e.g., Mercedes) while others seem to suffer from the data overflow (e.g., McLaren). This suggests that instead of collecting all possible data, the F1 teams should concentrate on collecting only relevant data and concentrate on development of effective predictive models.

Commented [GP15]: This is interesting result, but the implications could have been explained better. For example, how much value does investment into new analytical models deliver?

7. Limitations

My project has some limitations which represent opportunities for future investigations. First, the empirical setting analysed is highly dynamic and more traditional industries may not experience the same opportunities or capacity to deal with similar amounts of information and types of Big Data. However, my results could be useful for conventional industries that eventually will be confronted by Industry 4.0 Big Data revolution; it could help them to anticipate what to expect from the new regulation and the consequences of introducing radical innovation in a competitive market.

References [not included in the word count]

Ahlstrom D. 2010. Innovation and growth: How business contributes to society. *Academy of Management Perspectives* **24**(3): 11–24.

Akhtar P, Khan Z, Frynas JG, Tse YK, Rao-Nicholson R. 2017. Essential Micro-foundations for Contemporary Business Operations: Top Management Tangible Competencies, Relationship-based Business Networks and Environmental Sustainability. *British Journal of Management* : 1–20.

Armstrong SJ, Priola V. 2001. Individual Differences in Cognitive Style and their Effects on Task and Social Orientations of Self-Managed Work Teams. *Small Group Research* **32**(3): 283–312.

Aversa P, Cbantous L, Haefliger S. 2016. Decision-making with big data: How to design information systems with the practice in mind? *Working Paper* (January).

Aversa P, Furnari S, Haefliger S. 2015. Business model configurations and performance : A qualitative comparative analysis in Formula One racing , 2005 – 2013. *Industrial and Corporate Change* **24**(3): 655–676.

Bi F. 2014. How Formula One Teams Are Using Big Data To Get The Inside Edge. *Forbes.com* : 10.

Boyd D, Crawford K. 2012. Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information Communication and Society* **15**(5): 662–679.

Braganza A, Brooks L, Nepelski D, Ali M, Moro R. 2017. Resource management in big data initiatives: Processes and dynamic capabilities. *Journal of Business Research*. The Authors **70**: 328–337.

Brambor T, Clark WR, Golder M. 2006. Understanding interaction models: improving empirical analyses. *Political Analysis* **14**(1): 63–82.

Brockman BK, Morgan RM. 2003. The role of existing knowledge in new product innovativeness and performance. *Decision Sciences* **34**(2): 385–419.

Butler RJ, Carney M. 1986. Strategy and strategic choice: The case of telecommunications. *Strategic Management Journal* **7**(2): 161–177.

Cassiman B, Veugelers R. 2006. In search of complementarity in innovation strategy: Internal R&D and external knowledge acquisition. *Management Science* **52**(1): 68–82.

Chang Y-C, Chang H-T, Chic Bh-R, Chen M-H, Denge L-L. 2012. How do established firms improve radical innovation performance? The organizational capabilities view. *Technovation* **32**(7): 441–451.

Chari S, Katsikeas CS, Balabanis G, Robson MJ. 2014. Emergent Marketing Strategies and Performance: The Effects of Market Uncertainty and Strategic Feedback Systems. *British Journal of Management* **25**(2): 145–165.

Chen H, Chiang RHL, Storey VC. 2012. Business Intelligence and Analytics: From Big Data To Big Impact. *Mis Quarterly* **36**(4): 1165–1188.

Cialdini RB. 2001. Harnessing the science of persuasion. *Harvard Business Review* **79**(9): 72–81.

CNBC. 2014a. Spotlight on IT and data [Video file]. *www.cnbc.com/video/2014/05/28/spotlight-on-it-and-data.html*.

CNBC. 2014b. The history: IT and data [Video file]. *www.cnbc.com/video/2014/05/28/spotlight-on-it-and-data.html*.

Cohen J, Cohen P, West S, Aiken A. 2003. *Applied Multiple Regression/ Correlation Analysis for the Behavior Sciences.* Lawrence Erlbaum Associates: Mahwah, NJ.

Cohen WM, Levinthal DA. 1990. Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly*, Technology, organizations, and innovation, Starbuck WH, Whalen PS (eds). Cornell University, Johnson Graduate School **35**(1): 128–152.

Constantiou ID, Kallinikos J. 2015. New games, new rules: Big data and the changing context of strategy. *Journal of Information Technology* **30**(1): 44–57.
Covin JG, Prescott JE, Slevin DP. 1990. The effects of technological sophistication on strategic profiles, structure and firm performance. *Journal of Management Studies*. Blackwell Publishing Ltd **27**(5): 485–510.
Cyert RM, March JG. 1963. *A behavioral theory of the firm*. Englewood Cliffs.
Daft RL, Lengel RH. 1986. Organizational information requirements, media richness and structural design. *Management Science* **32**(5): 554–571.
Daft RL, Weick KE. 1984. Toward a Model of Organizations as Interpretation Systems. *Academy of Management Review*. JSTOR **9**(2): 284–295.
Danneels E, Kleinschmidt EJ. 2001. Product innovativeness from the firm's perspective: Its dimensions and their relation with project selection and performance. *Journal of Product Innovation Management* **18**(6): 357–373.
Davenport TH, Barth P, Bean R. 2012. How 'Big Data' Is Different. *MIT Sloan Management Review*, **54**(1): 43–46.
Davenport TH, Kudyba S. 2016. Designing and Developing Analytics-Based Data Products. *MIT Sloan Management Review* **58**(1): 58–63.
Duggan W. 2012. *Creative Strategy: A Guide for Innovation*. Columbia University Press: New York.
Dyck B, Starke FA, Mischke GA, Mauws M. 2005. Learning to build a car: An empirical investigation of organizational learning. *Journal of Management Studies*.
Edmondson AC. 2002. The Local and Variegated Nature of Learning in Organizations: A Group-Level Perspective. *Organization Science* **13**(2): 128–146.
Eisenhardt KM, Martin JA. 2000. Dynamic capabilities: What are they? *Strategic Management Journal*. John Wiley & Sons **21**(10–11): 1105–1121.
Fenn, J., LeHong, H. (2011). *Hype Cycle for Emerging Technologies*, 2011. Gartner, July.
Frizzo-Barker J, Chow-White PA, Mozafari M, Ha D. 2016. An empirical study of the rise of big data in business scholarship. *International Journal of Information Management* **36**(6): 403–413.
Gandomi A, Haider M. 2015. Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management* **35**(2): 137–144.
George G, Haas M, Pentland A. 2014. Big Data and Management. *Academy of Management Journal* **57**(2): 321–326.
Gino F, Pisano GP. 2011. Why Leaders Don ' t Learn From Success Why Leaders Don ' t Learn. *Harvard business review* **89**(4): 68–74.
Gitlin JM. 2017. Citrix isn't just for telecommuting, Red Bull Racing uses it at the track | Ars Technica. *Ars Technica*.
Goes P. 2004. Editor's Comments: Big Data and IS Research. *Mis Quarterly* **38**(3): 3–8.
Goodall AH, Pogrebna G. 2015. Expert leaders in a fast moving environment. *Leadership Quartely* **26**(2): 123–142.
Goodyear R, Agarwala-Rogers R. 1975. *Communication in Organizations*.
Hambrick DC, Abrahamson E. 2017. Assessing Managerial Discretion across Industries: A Multimethod Approach. *The Academy of Management Journal* **38**(5): 1427–1441.
Harrison, G. W., & List, J. A. (2004). Field experiments. *Journal of Economic Literature*, **42**(4): 1009-1055.
Helfat CE, Campo-Rembado MA. 2016. Integrative capabilities, vertical integration, and innovation over successive technology lifecycles. *Organization Science* **27**(2): 249–264.
Hodgkinson GP, Sparrow PR. 2002. *The competent organization: A psychological analysis of the strategic management process.* Open University Press: Buckingha, UK.
Hoisl K, Gruber M, Conti A. 2017. R&D team diversity and performance in hypercompetitive environments. *Strategic Management Journal* **38**(7): 1455–1477.
Hughes M, Hughes P, Morgan RE. 2007. Exploitative learning and entrepreneurial orientation

alignment in emerging young firms: Implications for market and response performance. *British Journal of Management* **18**(4): 359–375.
Hughes M, Morgan RE, Ireland RD, Hughes P. 2014, September 1. Social capital and learning advantages: A problem of absorptive capacity. *Strategic Entrepreneurship Journal* **8**(3): 214–233.
Ignatius J, Leen JYA, Ramayah T, Kah Hin C, Jantan M. 2012. The impact of technological learning on NPD outcomes: The moderating effect of project complexity. *Technovation*. Elsevier **32**(7–8): 452–463.
Jansen JJP, Van Den Bosch FAJ, Volberda HW. 2005. Managing Potential and Realized Absorptive Capacity: How Do Organizational Antecedents Matter? *The Academy of Management Journal* **48**(6): 999–1015.
Jenkins M. 2004. Innovative management: superior performance in changing times. *European Business Journal* **16**(1): 10–19.
Jenkins M. 2010. Technological Discontinuities and Competitive Advantage: A Historical Perspective on Formula 1 Motor Racing 1950-2006. *Journal of Management Studies* **47**(5): 884–910.
Jenkins M, Pasternak K, West R. 2007. *Performance at the limit: Business lessons from Formula 1 motor racing*. Cambridge: Cambridge University Press.
Kache F, Seuring S. 2017. Challenges and opportunities of digital information at the intersection of Big Data Analytics and supply chain management. *International Journal of Operations & Production Management* **37**(1): 10–36.
Kale P, Anand J. 2006. The Decline of Emerging Economy Joint Ventures: The Case of India. *California Management Review* **48**(3): 62–76.
Kearney C, Soleimanof S, Wales WJ. 2017. Examining Facilitative Configurations of Entrepreneurially Oriented Growth: An Information Processing Perspective. *British Journal of Management* **0**: 1–20.
Khanna T, Kartik V, Lane D. 2003. Formula One motor racing. *Harvard Business School Case* **September**: 403–412.
Lane PJ, Koka BR, Pathak S. 2006. The reification of absorptive capacity: A critical review and rejuvenation of the construct. *Academy of Management Review*. Academy of Management **31**(4): 833–863.
Leonard-Barton D. 1992. Core capabilities and core rigities: a paradox in managing new product development. *Strategic Management Journal* **13**: 111–125.
Lewin AY, Long CP, Carroll TN. 1999. The coevolution of new organizational forms. *Organization Science* **10**(5): 535–550.
Lichtenthaler U. 2009. Absorptive Capacity, Environmental Turbulence, and the Complementarity of Organizational Learning Processes. *Academy of Management Journal*. Academy of Management **52**(4): 822–846.
Linden, A., Fenn, J. (2003). Understanding Gartner's hype cycles. *Strategic Analysis Report N° R-20-1971*. Gartner, Inc.
Loebbecke C, Picot A. 2015. Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda. *Journal of Strategic Information Systems* **24**(3): 149–157.
Manovich L. 2011. Trending: The Promises and the Challenges of Big Social Data. *Debates in the Digital Humanities* : 460–475.
Marino A *et al.* 2015. Driving Performance via Exploration in Changing Environments: Evidence from Formula One Racing. *Organization Science* (August).
Markus LM. 2015. New games, new rules, new scoreboards: The potential consequences of big data. *Journal of Information Technology*.
Marr B. 2015. Big Data In Fast Cars -- How F1 And NASCAR Compete On Analytics. *Forbes*.
Matthias O, Fouweather I, Gregory I, Vernon A. 2017. Making sense of Big Data – can it transform operations management? *International Journal of Operations & Production*

Management. Emerald Publishing Limited **37**(1): 37–55.
Mayer-Schönberger V, Cukier K. 2013. *Big data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt.
McAfee A, Brynjolfsson E. 2012. Big Data. The management revolution. *Harvard Business Review* **90**(10): 61–68.
McKinsey Global Institute. 2011. Big data: The next frontier for innovation, competition, and productivity. *McKinsey Global Institute* (June): 156.
McNally RC, Cavusgil E, Calantone RJ. 2010. Product innovativeness dimensions and their relationships with product advantage, product financial performance, and project protocol. In *Journal of Product Innovation Management*, 27: 991–1006.
Moorman C, Miner AS. 1997. The impact of organizational memory on new product performance and creativity. *Journal of Marketing Research* (1): 91–160.
SAO/NASA Astrophysics Data System (1987) Telemetry: summary of concept and rationale, *NASA STI/Recon Technical Report* **89**, accessed via http://adsabs.harvard.edu/abs/1987STIN...8913455. on 29/07/2018
Nonaka T. 1994. A dynamic theory of organizational knowledge creation. *Organization Science* **5**(1): 14–37.
O'Reilly C, Tushman ML. 2004. The ambidextrous organization. *Harvard Business Review* **82**(April): 74–81.
Reinmoeller P, Ansari S. 2016. The Persistence of a Stigmatized Practice: A Study of Competitive Intelligence. *British Journal of Management* **27**(1): 116–142.
Rogers PR, Miller A, Judge WQ. 1999. Using information-processing theory to understand planning/performance relationships in the context of strategy. *Strategic Management Journal*. John Wiley & Sons, Ltd. **20**(6): 567–577.
Sakka O, Barki H, Côté L. 2016. Relationship between the interactive use of control systems and the project performance: The moderating effect of uncertainty and equivocality. *International Journal of Project Management*. Elsevier Ltd and Association for Project Management and the International Project Management Association **34**(3): 508–522.
Saldanha T, Mithas S, Krishnan MS. 2017. Leveraging Customer Involvement for Fueling Innovation: The Role of Relational and Analytical Information Processing Capabilities. *Management Information Systems Quarterly* **41**(1). Available at: https://aisel.aisnet.org/misq/vol41/iss1/16.
Seebode D, Jeanrenaud S, Bessant J. 2012. Managing innovation for sustainability. *R&D Management* **42**(3): 195–206.
Sen D, Ozturk M, Vayvay O. 2016. An Overview of Big Data for Growth in SMEs. *Procedia - Social and Behavioral Sciences*. Elsevier B.V. **235**(October): 159–167.
Sivarajah U, Kamal MM, Irani Z, Weerakkody V. 2017. Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research*. The Authors **70**: 263–286.
Smith WK, Tushman ML. 2005. Managing strategic contradictions: A top management model for managing innovation streams. *Organization Science* **16**(5): 522–536.
Spyropoulou S, Katsikeas CS, Skarmeas D, Morgan NA. 2017, February 10. Strategic goal accomplishment in export ventures: the role of capabilities, knowledge, and environment. *Journal of the Academy of Marketing Science*. Springer US : 1–21.
Stewart LA. 2010. The impact of regulation on innovation in the United States: A cross-industry literature review. *Commissioned by the Institute of Medicine Committee on Patient Safety and Health IT*.
Sylt C, Reid C. 2011. *Sylt, C., & Reid, C. 2011. Formula money. Southwick, Sussex: Wyndeham Grange.Formula money.* Southwick, Sussex: Wyndeham Grange.
Toet, W. (2013). Aerodynamics and aerodynamic research in Formula 1. *Aeronautical Journal*, **117**(1187), 1-26.
Trautmann G, Turkulainen V, Hartmann E, Bals L. 2009. Integration in the global sourcing organization-An information processing perspective. *Journal of Supply Chain Management*

45(2): 57–74.
Tushman ML, Nadler DA. 1978. Information Processing as an Integrating Concept in Organizational Design. *Academy of Management Review*. Academy of Management **3**(3): 613–624.
de Visser M, Faems D, Visscher K, De Weerd-Nederhof P. 2014. The impact of team cognitive styles on performance of radical and incremental NPD projects. *Journal of Product Innovation Management* **31**(6): 1167–1180.
Wanga S, Wana J, Zhang D, Li D, Zhanga C. 2016. Towards smart factory for industry 4.0: A self-organized multi-agent system with big data based feedback and coordination. *Computer Networks* **101**(1): 158–168.
Westervelt A. 2012. How international regulations are changing American supply chains. *Forbes* **May**.
Wiseman RM, Catanach C. 1997. A longitudinal disaggregation of operational risk under changing regulations: Evidence from the savings and loan industry. *The Academy of Management Journal* **40**(4): 799–830.
Wu Y, Cegielski CG, Hazen BT, Hall DJ. 2013. Cloud computing in support of supply chain information system infrastructure: Understanding when to go to the cloud. *Journal of Supply Chain Management* **49**(3): 25–41.