Divining Insights: Visual Analytics Through Cartomancy

Andrew McNutt
University of Chicago
Chicago, IL 60637, USA
mcnutt@uchicago.edu

Michael Correll
Tableau Research
Seattle, WA 98103
mcorrell@tableau.com

Anamaria Crisan
Tableau Research
Seattle, WA 98103
acrisan@tableau.com

CHI Conference on Human Factors in Computing Systems Extended Abstracts, April 25–30, 2020, Honolulu, HI, USA
ACM 978-1-4503-6819-3/20/04.
https://doi.org/10.1145/3334480.3381814

Abstract

Our interactions with data, visual analytics included, are increasingly shaped by automated or algorithmic systems. An open question is how to give analysts the tools to interpret these “automatic insights” while also inculcating critical engagement with algorithmic analysis. We present a system, *Sortilège*, that uses the metaphor of a Tarot card reading to provide an overview of automatically detected patterns in data in a way that is meant to encourage critique, reflection, and healthy skepticism.

Author Keywords

Visual analytics; information visualization; automated insights; divination

CCS Concepts

•Human-centered computing $\rightarrow$ Visualization systems and tools;

Introduction

An unsolved challenge in visual analytics is how to strike the proper balance between automated and manual exploration of data [22]. Automated methods based on statistical modelling and machine learning can assist the analyst in sorting through large amounts of complex data, focus attention on what is important, and provide guidance for users without experience or expertise in analytics. These

Figure 1: *Cartomancy* is the use of playing cards (such as Tarot cards) for divination. Tarot has a long history, and an a correspondingly sprawling collection of variations in detail, form, and interpretation. The Rider-Waite deck [49] is a common set of Tarot cards on which we base the form and interpretation of the cards in this work.

methods, if taken to the extreme, promise the possibility of automatically generating "insights" from datasets [17, 43, 50]: entirely automating the potentially tedious process of exploratory data analysis. This promise comes with several dangers. Automated analysis methods can be *opaque and inscrutable*: what the system is doing may be unclear or unverifiable. Automated methods can be *brittle and biased*: these systems, by construction, lack context, history, or plain common sense about what things in the dataset are important. In the worst case, they can function as "p-hacking machines" [37] that highlight chance occurrences in the dataset that are unlikely to generalize or replicate [53]. Automated methods can promote a lack of *skepticism or critical thought*: they often fail to communicate uncertainty about their findings [24] while at the same time appearing to present unimpeachable facts. Lastly, automated analytical methods can *discourage exploration or discovery* by presenting a static list of insights, usually in some rank order, and so potentially avoiding analysis beyond the obvious, high-level trends in the data.

As a critical design exercise [2], we present *Sortilège*, an "automated insights" system that uses metaphors and processes taken from Tarot to surface potentially relevant patterns in a dataset. A demo of *Sortilège* is deployed at https://vis-tarot.netlify.com/. The code for the application is available at https://osf.io/pwbmj/ under an MIT license.

Like in traditional Tarot (Figure 1), in *Sortilège*, the user mentally formulates a question and then consults a spread of cards. Each space for a card has a particular divinatory meaning (for instance, in a three card spread, the three cards may have meanings about the context, hidden influences, and provide advice about a situation, respectively). The user then deals cards into the spread. Each card has an associated set of meanings or connotations that the user interprets in light of their question and the card's location (for instance, "The Tower" is traditionally associated with an unpredictable and potentially disastrous change). *Sortilège*'s cards are a combination of hints, guidance, and questions about data analytics (forming our Major Arcana) and visualizations of potentially relevant patterns in a particular dataset (forming our Minor Arcana). We intend for *Sortilège* to act as an example of a mixed-initiative system of automated analysis that encourages reflection, discovery, and chance: guidance for a later exploratory or confirmatory data analysis session, rather than an exhaustive, immutable, and authoritative report of "insights."

Sortilège foregrounds many of our concerns with automated analytics: it is largely inscrutable, fragile, and operates in many ways like the "black box" systems we criticize. And yet, we argue that *Sortilège* is in some ways "safer" than existing designs. It encourages critical thought and skepticism, makes no pretense towards certainty, requires users perform the effort of interpretation, and supports and encourages "serendipity strategies" [28] (such as varying routines and searches for patterns) that could result in better analytical outcomes than a static list of "insights."

Background

Commercial data analytics software has begun to employ statistical modeling or machine learning to automatically generate "insights" about particular datasets. For instance, PowerBI's "Quick Insights" panel creates charts of potentially relevant features in a dataset such as correlated fields or outliers [36]. Similarly, Tableau's "Explain Data" feature creates secondary explanatory charts that attempt to account for a particular unexpected value in a chart [42].

From within academia, there has been a growing set of examples of systems that attempt to perform some combination of *augmenting* exploration of datasets with automatically generated charts [43], *recommending* potentially interesting charts in order to jumpstart an exploratory data analysis session [52], or even entirely *automating* the process of visualizing noteworthy facts or patterns in a dataset [50] (amongst many other potential tasks; see Ceneda et al. [9] and Collins et al. [10] for an overview).

The appeal of these automated or mixed-initiative approaches is obvious: they can dramatically reduce the expenditure of time or the required statistical expertise necessary to discover an important insight in the data. Even if the resulting automatically generated charts are not immediately insightful, they can provide useful starting points or shortcuts for deeper analysis. They can also function as safeguards or sanity checks [12], alerting analysts to facts about their dataset that they might have missed. And yet, these methods have many potential shortcomings and dangers.

Critiques of Automated Analytics
Heer [22] identifies designing for a useful mixture of automation and agency as an emerging challenge in HCI, visual analytics (VA) included. Automated methods and recommendation systems often fall short in this respect. We highlight here four ways in which automated analytics methods and other "insight generation" tools can be problematic.

Opaque
The methods by which charts are automatically generated are often statistical "black boxes" that defy easy interpretation. For instance, both Data2Vis [15] and VizML [23] use neural nets to supply recommendations. Making neural networks interpretable (or even what it *means* for such models to be interpretable) is an open problem in machine learning [27]. In the absence of information, the user may blindly trust that system is offering good suggestions [29]. Our increasing reliance on these uninterpretable (and so un-auditable and unappealable) systems contributes to systematic injustice and societal inequality [32, 33].

Inflexible
Automated analytics systems are often meant to be domain-agnostic, which means that they cannot adapt to the peculiarities of particular domains. Since the very definition of an "insight" is complex and multifaceted [31], most systems only capture a small subset of such insights, using a hodgepodge of unrelated algorithms that may or may not match human intuitions. For instance, many methods for automatically detecting outliers rely on distance from a central tendency, and so may miss "interior" outliers that humans readily notice [51].

This inflexibility extends to visualization types as well. Automated analytics systems can only present the types of charts they have been designed to generate. For instance, Draco [30] is unable to suggest a pie chart (even in contexts where one might be appropriate) because it operates over Vega-Lite [39] which can not currently describe a pie chart. Many analytics systems present only a core set of simplistic charts, despite the expansive language of charts and graphs for highlighting important aspects of data.

Brittle
Automated analytics systems make exhaustive searches of the data, surfacing charts that have been chosen in such a way as to maximize a particular interest (such as surprise [46]). Without proper controls for false positives or other aspects related to the multiple comparisons problem, these systems can act as "p-hacking machines" [37], surfacing "insights" of dubious quality [4]. Many, if not most, insights surfaced from the unprincipled visual exploration of the data may fail to generalize [53]. By taking all or most analytical paths simultaneously, the validation of any particular insight is difficult [18, 37]. Even if analyzed responsibly, many VA

systems fail to communicate uncertainty information [24], giving analysts little in the way of guidance for determining the strength or certainty surrounding what they have found.

Domineering

Charts are inherently rhetorical devices and carry a heavy rhetorical weight [25]. This often inspires users to trust charts as true pictures of objective data [11, 19, 35]. When such charts are presented in the context of recommendations from "smart" statistical procedures, this can strengthen the rhetorical impression of authority and veracity. This unearned impression of objectivity and expertise—this privileged "view from nowhere" [3]—can exacerbate these weaknesses, as analysts may fail to properly vet charts presented to them from a trusted recommender.

Similarly, automated analytics and recommendation systems may fail to adapt to the analytical needs of the user, and enforce a particular path through the data. Exploratory data analysis entails the ability to flexibly, iteratively, and occasionally serendipitously [45] explore the data. By presenting the same "starting places" and the same suggested analyses each time, automated systems stifle this freedom.

Folk and Occult Algorithms

Models and algorithms can be difficult to understand. They are often inherently complex, require specialized or esoteric knowledge to build, or are intentionally obfuscated through appeals to trade secrets or minimizing the complexity of the end-user experience. In the absence of a detailed understanding of an algorithm, "folk theories" of algorithmic performance arise [14]. These folk understandings extend also to peoples' understanding of data visualizations [35], in which skepticism concerning the source or veracity of data can be swept aside by the sheer rhetorical force of a chart as an objective representation of truth [25].

With similar motivations to our work, Browne & Swift [8] construct a neural net whose outputs are interpreted via a ritual séance, playing with the dual meanings of occult (as in hidden from explanation or examination, or as in mystical or supernatural). Similarly, Lee et al. [26] use sibylline obfuscation and occult symbology to present the results of discourse analysis algorithms. While we do not deny that *Sortilège* shares these projects' goal of drawing critical attention to the fact that many of the algorithms in which we have invested a great deal of political, social, and economic power are inexplicable to large audiences, we adopt occult practices here for more positive reasons as well. Sultana et al. [44] remark that integrating occult practices into HCI can function as a way of widening the scope of our audience, and promoting re-examination of the epistemologies and power structures that underlie our design work.

Our work bares a close resemblance to prior work on furthering creative projects. Eno's Oblique Strategies [20] makes use of the randomness latent in a deck of ambiguous-but-provocative prompts as a mechanism for re-framing creative projects. Börütecene & Buruk [5] make use of a Ouija board as a way to enrich the design process. Sengers et al. [41] push designers to promote skepticism and reflection in their users through the form of their designs.

Our choice to rely on the trappings of Tarot (with its elements of randomness, personal interpretations, and intentional ambiguity) locates our work within an emerging trend that pushes back against traditional views of what a visualization system should support. For instance, Bradley et al. [6] call for a "slow analytics" movement that may not present answers as quickly as possible, but has the benefit of increased human engagement and ownership over the analysis process. Thudt et al. [45] and Alexander et al. [1] tout the benefits "serendipitous" visualization tools

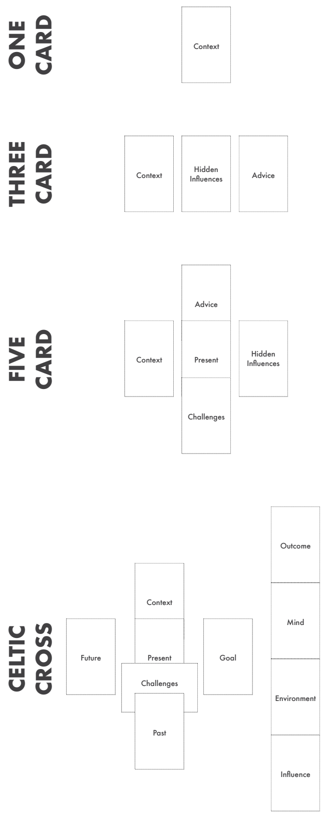

Figure 2: The available card spreads in *Sortilège* (shown here) are drawn from traditional Tarot practices. Future development of our system might include additional spreads specifically tuned to the task of visual analytics, such as an analytical visionary tableau.

that encourage potentially aimless or unguided exploration of datasets. Tarot, with its heavy reliance on human-driven interpretation of stochastic systems, is in accord with both of these emerging design directions for visual analytics.

Sortilège

Sortilège is a web-based skeuomorphic prototype that is intended to mimic the act of using Tarot cards for divination, but tuned to the process of investigating important properties of a particular dataset as part of an initial step prior to more in-depth visual analytics.

The user (hereafter the *querent*) uploads or selects a pre-loaded dataset and selects a Tarot *spread*. In Tarot, spreads are spatial arrangements of cards where each location has a particular divinatory meaning. Some of these spreads are simple (for instance, a common three card spread arranges the cards in a row with a spaces for cards relating to the “Past,” “Present,” and “Future” respectively). Others, like the Celtic Cross, are complex arrangements of many cards. In some cases the customary meanings of the locations do not neatly correspond to the intended task of querying a *dataset*, as opposed to introspection or divination concerning a *person* or *situation*. In those cases we have changed the labels and context of those locations in a way we believe preserves the general spirit of the spread. We depict the spreads available in *Sortilège* in Figure 2.

Once the querent has chosen a spread, *Sortilège* generates (and shuffles) a “deck” of insights relevant to the data. A Tarot deck is composed of two types of cards: *Major Arcana*, which are often strong and potent omens, and *Minor Arcana*, divided into four suits, which often represent less important or salient forces. The querent clicks on the deck of cards to deal one card at a time into the spread, such that they reflect on that card’s meaning both in its particular location in the spread, as well as in the context of the other cards that have been uncovered so far. Figure 3 shows an example of the system after completing a reading. In the following sections we describe how the cards in our deck are generated.

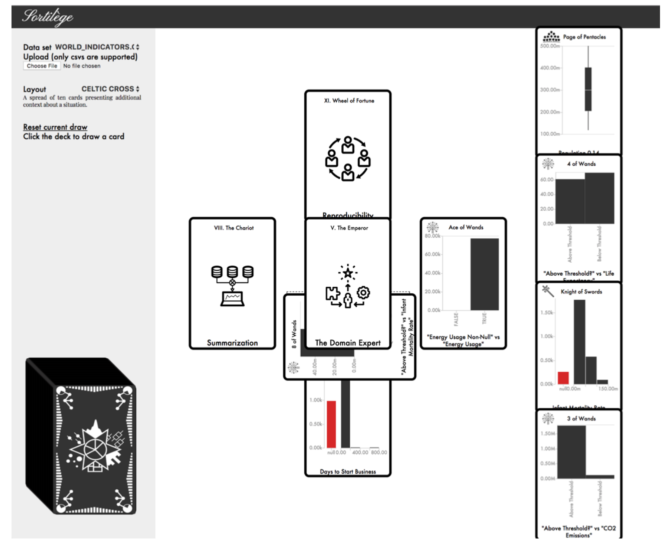

Figure 3: A querent seeks knowledge of the World Indicators dataset [21]. They’ve drawn into the Celtic cross layout. A mixture of advice (in the form of Major Arcana) and potentially interesting patterns (in the form of Minor Arcana) suggest further routes of exploration with this dataset.

Major Arcana

In traditional Tarot, the twenty-two cards of the Major Arcana represent iconic forces with a multiplicity of meanings. For instance, “The Star” (the 17th card of the Major Arcana) may indicate contentment, hope, inspiration, or opportunity.

Our Major Arcana promote skepticism and reflection, both

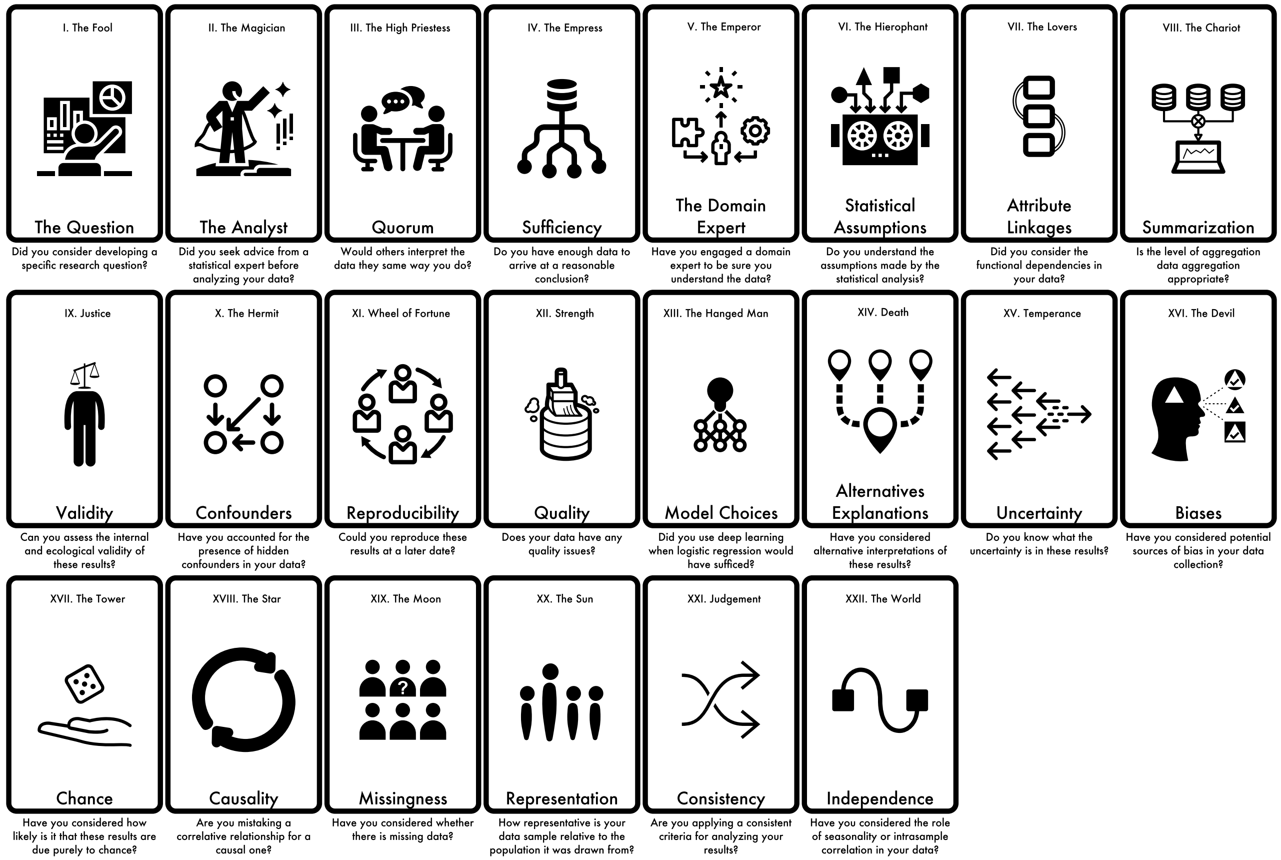

Figure 4: The cards constituting the Major Arcana in *Sortilège*. Each card provides a thematic interpretation of the traditional meaning of card consisting of a new name, an illustrative design, and a question designed to foster introspection and skepticism about the analysis process.

on any particular insight received from the data, but also in general on how the dataset is organized, collected, and used. Drawing on both common pitfalls in visual analytics (such as those laid out by Bresciani & Eppler [7]), as well as ethical "checklists" for data science (such as those proposed by Patil et al. [34]), we created a set of potential concerns or forces that may be at play when analyzing data. As with similar efforts, such as the Control-Alt-Hack card

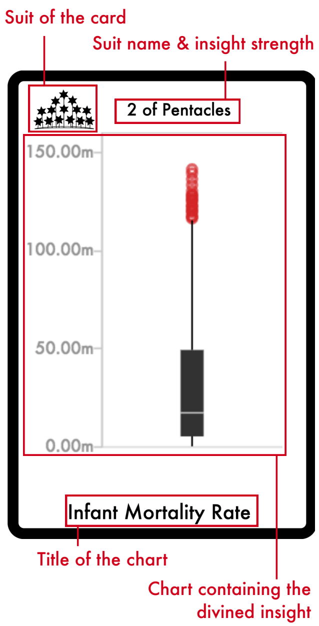

Figure 5: An annotated card drawn from the Minor Arcana, here depicting the volume of outliers in the World Indicators dataset [21]. In this case, the *suit* of the card, Pentacles, indicates that this is a field with at least one extreme value. But the relatively small *rank* of the card indicates that other fields in the dataset may have even more extreme outliers.

game [13], it is our hope that these questions and positions are broad enough to be useful in many scenarios, but to also encourage creative thought about how they might apply to the specific situation in mind.

Figure 4 shows all of our Major Arcana. These cards ask direct questions to the querent following four themes of visual analyses: the people involved in the analysis, the data as an artifact, issues arising from models of the data, and issues arising from the analysis process itself. To wit, "The Sun" (the 20th card of the Major Arcana) asks the querent who is *counted*, while "The Emperor" (the 5th card) prompts them to consider who is doing the *counting* [11, 16].

Minor Arcana

In traditional Tarot, the Minor Arcana are divided into four suits. For instance, in the Rider-Waite deck [49] and many others decks, these suits are Wands, Cups, Pentacles, and Swords. Within each suit the cards have numbers, from the Ace to the Page, Queen, and King of that suit in increasing order of “strength.” As with the Major Arcana, each card has different customary meanings, with the difference that each suit is often thought of as having thematic divinatory properties in common. For instance, traditional Sword cards are often associated with problems, conflicts, and power.

In *Sortilège*, each suit is associated with a general class of statistical patterns or properties, each of which are automatically generated. Modeled on other recommendation systems like Voyager [52], SeeDB [47], and Draco [30], these recommendations are presented on the face of the cards as Vega [40] or Vega-Lite [39] charts. Figure 5 shows an annotated example of a Minor Arcana card.

The types of “insights” we generate will vary according to the four suits of the Minor Arcana: Wands, Cups, Pentacles, and Swords. *Sortilège* generates univariate (Swords, Pentacles) and bivariate (Wands, Cups) charts from fields in the dataset, in increasing order of insight “strength.” For instance, the suit of Swords presents insights pertaining to data quality within a field; the King of Swords indicates the field with the largest number of null or invalid values, and the Queen the field with the second largest, until either the number of fields is exhausted or we have reached the ace of that particular suit, whichever happens first. This results in a maximum of 52 potential insights about the data, although this number can be lower depending on the size and complexity of the dataset.

We now detail how *Sortilège* generates each of the Minor Arcana, although we note that the algorithms we use to propose “insights” are intentionally simplistic, and have many of the drawbacks we enumerate earlier in the paper. Our goal with this critical design exercise was to explore if automated insights could be presented to the user in safer and more serendipitous ways, rather than implement a deep and statistically complex recommendation engine. Figure 6 shows the Minor Arcana created for a particular dataset.

The Suit of Wands contains patterns related to categorical variance. That is, large swings in values across a field (say, very high sales in one region of the country, and low sales in another) would be very strong insights for this suit. *Sortilège* searches every combination of quantitative variable, categorical variable, and two aggregation functions (*sum* or *mean*) and calculates the variance normalized across all categories. Higher variance yields a higher strength score. We show these insights on each of the cards as a categorical bar chart of the relevant fields.

The Suit of Cups contains patterns related to correlations and relationships between fields. *Sortilège* computes the Pearson’s correlation coefficient between every combination of numeric fields. The higher this coefficient, the

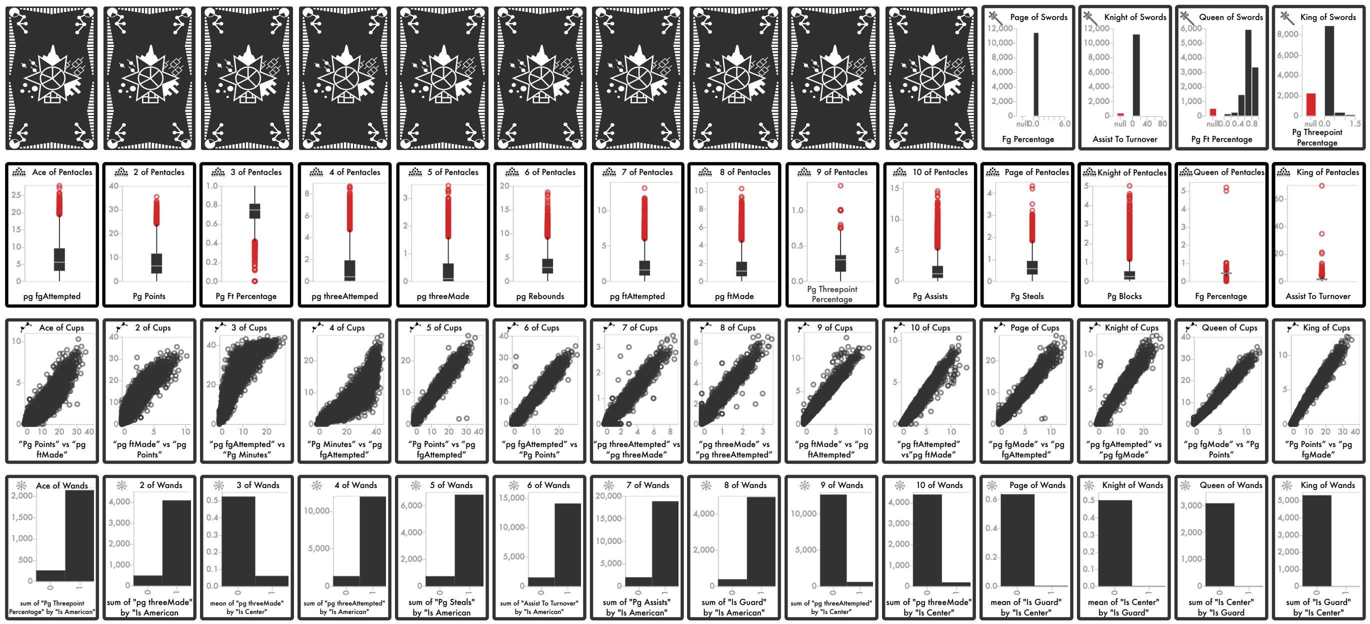

Figure 6: The cards of the Minor Arcana for a basketball statistics dataset. *Sortilège* only generate cards appropriate to the phenomenon that each suit describes. For instance, as there were only four fields with missing data, we generate only four cards in the suit of Swords; the rest of the suit is indicated here as face down cards.

higher the strength score (although we exclude fields with perfect correlation, which are often, but not always, trivial re-codings or dependencies in the data). Each card in this suit depicts a scatterplot of the two relevant fields.

The Suit of Pentacles contains patterns related to extreme or unexpected values within numeric fields. *Sortilège* computes the z-score of each value in each numeric field. The higher the maximum absolute value of z score in a particular field, the higher the strength score. Each member of this suit shows a boxplot of the relevant field.

The Suit of Swords contains patterns related to data quality concerns. *Sortilège* computes the proportions of rows in each field that are null, empty, or undefined. The higher the proportion of missing values, the higher the strength score. Members of this suit render these issues through a histogram of the relevant field.

Case Study

As a proof of concept, one of the authors consulted a simple three card spread in the context of the World Bank's "World Indicators" [21] dataset (Figure 7). This dataset is used as the backing data for many data advocacy groups, such as Rosling's GapMinder [38], and reports yearly country-level critical statistics stretching back a number of years.

The first card the querent draws is meant to connect to "Context": what sort of background information is important before performing analysis? The resulting card is the 8 of Swords (associated in traditional Tarot with powerlessness or immobility), showing the querent that a large percentage of data concerning CO_2 emissions is null, undefined, or otherwise missing from the dataset. This suggests that questions about emissions may be unanswerable or unaddressed (or at the very least the querent should be cautious about claims using this field). That this is only the *8* of Swords is troubling as well: it indicates that there are many other fields with even higher proportions of missing data.

The second card drawn by the querent is placed into the slot for "Hidden Influences," which deals with undercurrents in the present analysis. The resulting card is the Queen of Pentacles (associated in traditional Tarot with luxury and wealth), showing the querent that, while many countries have very few incoming tourists, there are some countries with extremely high numbers, reducing the central mass of "un-visited" countries to near invisibility in the box plot.

The last card, in the slot for "Advice," concerns recommendations for moving forward with the analysis or applying the lessons learned in the divination session. It is the Major Arcana of The Devil (in traditional Tarot, a sign of unbridled id or harmful attachment). Here its meaning is linked to the notions of biases (cognitive or otherwise) that may be at play when measuring or modeling data.

The querent had several questions and potential next steps as a result of this reading. For instance, how prevalent are the data quality issues observed in the divination in the

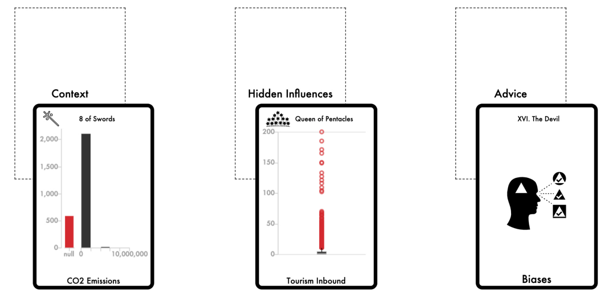

Figure 7: An example draw received by one of the paper authors, acting as querent, when consulting the World Indicators dataset.

rest of the dataset? Given the existence of a contingent of outlier countries for metrics like tourism, is it possible that there are other groups of high or low-metric countries that are skewing aggregate values across the rest of the dataset? Worse yet, if countries receive staggeringly different amount of attention, is it possible that different countries are more or less *measured* (both in this dataset, and in general)? If so, might this bias cause an analyst to under- or over-estimate the scale of global problems? Questions and reflections like these are of course idiosyncratic to the querent, but we believe that the ambiguous and polysemic nature of card readings can promote deep reflection.

This narrative represents just one particular draw from the deck connected to this dataset. A different querent will doubtless receive an entirely different set of cards, with entirely different meanings, that promote entirely different questions. The power of this stochastic and limited view is that it raises questions about what the querent *didn't* see, or *might have* seen in the data. For instance, if the Queen of Pentacles has this many outliers, what might the King of Pentacles look like? We contrast this uncertainty and caution with traditional automated insights systems, where the querent might be left with the impression that only the insights that made it to the panel or the recommendation screen are “important,” and the rest of the dataset largely matches their expectations.

Discussion

As with all critical design projects, we charge the reader to reflect on *Sortilège* as a provocation. For instance, what is safe or unsafe about automated analysis, and how can it be made safer? What responsibilities as designers do we have to transparently communicate to audiences with diverse areas of expertise? How can we design analytic systems to build up and build upon this expertise? How responsible are we as designers when people use the analytical systems we design come to erroneous conclusions?

The mystical black box we present with *Sortilège* is not fundamentally that much different from the technological black box presented by conventional VA systems. Both systems have an element of blind trust. However, we believe that conventional analytics systems go out of their way to build authority and perceived accuracy through tactics like suppressing uncertainty information [24] or using “AI washing” [48] to convey unearned expertise or complexity. They are occult systems in the original meaning of the term: hidden knowledge. One cannot shy away from this occult nature in *Sortilège*. Similarly, by forcing a burden of interpretation and confabulation onto the querent, *Sortilège* cuts off the easy excuse that generated insights are objective, neutral, or self-generating facts about the data. The human deals the cards, creates their interpretations, and links them together. They are an active, rather than passive, participant in the analysis process, random and arbitrary though it might be at its heart.

Sortilège's manifestation as a web app allows it the optimal flexibility to analyze datasets and mimics the digital-ephemerality of insights divined from conventional VA systems. Unlike those systems however, if a dataset were of particular significance or importance then a material copy of the corresponding deck could be physicalized and associated rituals of analysis engaged. Instead of monitoring a dashboard to assess key metrics, one could imagine performing a daily reading of a newly printed card deck.

Conclusion

We have described *Sortilège*, an automated visual insight system wrought in the medium of a Tarot deck. Our system allows users to explore their data through a traditional or familiar set of charts, alongside a collection of reflective prompts, that have been recontextualized and defamiliarized through an occult lens. Our goal in employing the occult is that users of this system are prompted by the form and design of our application to engage in the visual analysis process with a greater skepticism and unshackled curiosity. It is our hope that this critical reflection will carry over into analysts' everyday analysis practices. Further, we hope that they will refrain from blindly trusting visualization recommendation and automated analysis systems, that they will think more carefully about the who-how-and-why of their data, and, at the very least, question their assumptions when working with data.

Acknowledgements

We thank the divine hands of fate for having dealt us the cards that have allowed us to conduct this work, as well as our reviewers for their helpful readings and commentary.

REFERENCES

[1] Eric Alexander, Joe Kohlmann, Robin Valenza, Michael Witmore, and Michael Gleicher. 2014.

Serendip: Topic Model-Driven Visual Exploration of Text Corpora. In *2014 IEEE Conference on Visual Analytics Science and Technology (VAST)*. IEEE, 173–182. DOI: http://dx.doi.org/10.1109/VAST.2014.7042493

[2] Jeffrey Bardzell and Shaowen Bardzell. 2013. What is "Critical" about Critical Design?. In *Proceedings of the 2013 CHI Conference on Human Factors in Computing Systems*. ACM, 3297–3306. DOI: http://dx.doi.org/10.1145/2470654.2466451

[3] Shaowen Bardzell and Jeffrey Bardzell. 2011. Towards a Feminist HCI Methodology: Social Science, Feminism, and HCI. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 675–684. DOI:http://dx.doi.org/10.1145/1978942

[4] Carsten Binnig, Lorenzo De Stefani, Tim Kraska, Eli Upfal, Emanuel Zgraggen, and Zheguang Zhao. 2017. Toward Sustainable Insights, or Why Polygamy is Bad for You. In *CIDR 2017, 8th Biennial Conference on Innovative Data Systems Research, Chaminade, CA, USA, January 8-11, 2017, Online Proceedings*. www.cidrdb.org. http://cidrdb.org/cidr2017/papers/p56-binnig-cidr17.pdf

[5] Ahmet Börütecene and Oundefineduz 'Oz' Buruk. 2019. Otherworld: Ouija Board as a Resource for Design. In *Proceedings of the Halfway to the Future Symposium 2019 (HTTF 2019)*. Association for Computing Machinery, New York, NY, USA, Article Article 4, 4 pages. DOI: http://dx.doi.org/10.1145/3363384.3363388

[6] Adam James Bradley, Victor Sawal, and Christopher Collins. 2019. Approaching Humanities Questions Using Slow Visual Search Interfaces. IEEE 4th Workshop for Visualization and the Digital Humanities.

[7] Sabrina Bresciani and Martin J Eppler. 2015. The Pitfalls of Visual Representations: A Review and Classification of Common Errors Made While Designing and Interpreting Visualizations. *Sage Open* 5, 4 (2015). DOI: http://dx.doi.org/10.1177/2158244015611451

[8] Kieran Browne and Ben Swift. 2018. The Other Side: Algorithm as Ritual in Artificial Intelligence. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, alt11. https://doi.org/10.1145/3170427.3188404

[9] Davide Ceneda, Theresia Gschwandtner, Thorsten May, Silvia Miksch, Hans-Jörg Schulz, Marc Streit, and Christian Tominski. 2016. Characterizing Guidance in Visual Analytics. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (2016), 111–120. DOI: http://dx.doi.org/10.1109/TVCG.2016.2598468

[10] Christopher Collins, Natalia Andrienko, Tobias Schreck, Jing Yang, Jaegul Choo, Ulrich Engelke, Amit Jena, and Tim Dwyer. 2018. Guidance in the human–machine analytics process. *Visual Informatics* 2, 3 (2018), 166–180. DOI: http://dx.doi.org/10.1016/j.visinf.2018.09.003

[11] Michael Correll. 2019. Ethical Dimensions of Visualization Research. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, 188. DOI: http://dx.doi.org/10.1145/3290605.3300418

[12] Michael Correll, Mingwei Li, Gordon L. Kindlmann, and Carlos Scheidegger. 2019. Looks Good To Me: Visualizations As Sanity Checks. *IEEE Trans. Vis. Comput. Graph.* 25, 1 (2019), 830–839. DOI: http://dx.doi.org/10.1109/TVCG.2018.2864907

[13] Tamara Denning, Adam Lerner, Adam Shostack, and Tadayoshi Kohno. 2013. Control-Alt-Hack: The Design and Evaluation of a Card Game for Computer Security Awareness and Education. In *Proceedings of the 2013 ACM SIGSAC Conference on Computer & Communications Security (CCS '13)*. Association for Computing Machinery, New York, NY, USA, 915–928. DOI:http://dx.doi.org/10.1145/2508859.2516753

[14] Michael A DeVito, Jeffrey T Hancock, Megan French, Jeremy Birnholtz, Judd Antin, Karrie Karahalios, Stephanie Tong, and Irina Shklovski. 2018. The algorithm and the user: How can HCI use lay understandings of algorithmic systems?. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, panel04. DOI: http://dx.doi.org/10.1145/3170427.3188404

[15] Victor Dibia and Çağatay Demiralp. 2019. Data2Vis: Automatic Generation of Data Visualizations Using Sequence-to-Sequence Recurrent Neural Networks. *IEEE Transactions on Visualization and Computer Graphics* 39, 5 (2019), 33–46. DOI: http://dx.doi.org/10.1109/MCG.2019.2924636

[16] Catherine D'Ignazio and Lauren Klein. 2016. Feminist Data Visualization. In *IEEE VIS: Workshop on Visualization for the Digital Humanities (VIS4DH)*.

[17] Rui Ding, Shi Han, Yong Xu, Haidong Zhang, and Dongmei Zhang. 2019. QuickInsights: Quick and Automatic Discovery of Insights from Multi-Dimensional Data. In *Proceedings of the 2019 International Conference on Management of Data*. ACM, 317–332. DOI: http://dx.doi.org/10.1145/3299869.3314037

[18] Pierre Dragicevic, Yvonne Jansen, Abhraneel Sarma, Matthew Kay, and Fanny Chevalier. 2019. Increasing the Transparency of Research Papers with Explorable Multiverse Analyses. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, 65. DOI: http://dx.doi.org/10.1145/3290605.3300295

[19] Johanna Drucker. 2012. Humanistic theory and digital scholarship. *Debates in the digital humanities* (2012), 85–95.

[20] Brian Eno and Peter Schmidt. 1975. Oblique Strategies. *Opal.(Limited edition, boxed set of cards.)[rMAB]* (1975).

[21] World Bank Group. 2019. World Development Indicators. http://datatopics.worldbank.org/world-development-indicators/. (2019).

[22] Jeffrey Heer. 2019. Agency plus automation: Designing artificial intelligence into interactive systems. *Proceedings of the National Academy of Sciences* 116, 6 (2019), 1844–1850. DOI: http://dx.doi.org/10.1073/pnas.1807184115

[23] Kevin Hu, Michiel A Bakker, Stephen Li, Tim Kraska, and César Hidalgo. 2019. VizML: A Machine Learning Approach to Visualization Recommendation. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, 128. DOI: http://dx.doi.org/10.1145/3290605.3300358

[24] Jessica Hullman. 2020. Why Authors Don't Visualize Uncertainty. *IEEE transactions on visualization and computer graphics* 26, 1 (2020), 130–139. DOI: http://dx.doi.org/10.1109/TVCG.2019.2934287

[25] Helen Kennedy, Rosemary Lucy Hill, Giorgia Aiello, and William Allen. 2016. The work that visualisation conventions do. *Information, Communication & Society* 19, 6 (2016), 715–735. DOI: http://dx.doi.org/10.1080/1369118X.2016.1153126

[26] Joyce Lee, Sejal Popat, and Soravis Prakkamakul. 2019. Exploring the “Magic” of Algorithmic Predictions with Technology-Mediated Tarot Card Readings. https://www.ischool.berkeley.edu/projects/2019/exploring-magic-algorithmic-predictions, (2019). Accessed: 2020-01-03.

[27] Zachary C Lipton. 2018. The Mythos of Model Interpretability. *Commun. ACM* 61, 10 (2018), 36–43. DOI:http://dx.doi.org/10.1145/3233231

[28] Stephann Makri, Ann Blandford, Mel Woods, Sarah Sharples, and Deborah Maxwell. 2014. “Making my own luck”: Serendipity Strategies and Howto Support Them in Digital Information Environments. *Journal of the Association for Information Science and Technology* 65, 11 (2014), 2179–2194. DOI: http://dx.doi.org/10.1002/asi.23200

[29] Andrew McNutt and Gordon Kindlmann. 2018. Linting for Visualization: Towards a Practical Automated Visualization Guidance System. In *VisGuides: 2nd Workshop on the Creation, Curation, Critique and Conditioning of Principles and Guidelines in Visualization*.

[30] Dominik Moritz, Chenglong Wang, Greg L Nelson, Halden Lin, Adam M Smith, Bill Howe, and Jeffrey Heer. 2019. Formalizing Visualization Design Knowledge as Constraints: Actionable and Extensible Models in Draco. *IEEE Transactions on Visualization and Computer Graphics* 25, 1 (2019), 438–448. DOI: http://dx.doi.org/10.1109/TVCG.2018.2865240

[31] Chris North. 2006. Toward Measuring Visualization Insight. *IEEE computer graphics and applications* 26, 3 (2006), 6–9. DOI: http://dx.doi.org/10.1109/MCG.2006.70

[32] Cathy O’Neil. 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens democracy*. Broadway Books.

[33] Frank Pasquale. 2015. *The Black Box Society*. Harvard University Press.

[34] DJ Patil, Hilary Mason, and Mike Loukides. 2018. *Ethics and Data Science*. O’Reilly Media, Inc.

[35] Evan M Peck, Sofia E Ayuso, and Omar El-Etr. 2019. Data is Personal: Attitudes and Perceptions of Data Visualization in Rural Pennsylvania. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, 244. DOI: http://dx.doi.org/10.1145/3290605.3300474

[36] Microsoft PowerBI. 2019. Generate data insights automatically with Power BI. https://docs.microsoft.com/en-us/power-bi/service-insights. (2019).

[37] Xiaoying Pu and Matthew Kay. 2018. The Garden of Forking Paths in Visualization: A Design Space for Reliable Exploratory Visual Analytics: Position Paper. In *IEEE VIS: Evaluation and Beyond-Methodological Approaches for Visualization (BELIV)*. IEEE, 37–45. DOI: http://dx.doi.org/10.1109/BELIV.2018.8634103

[38] Hans Rosling. 2009. GapMinder Foundation. http://www.gapminder.org. (2009).

[39] Arvind Satyanarayan, Dominik Moritz, Kanit Wongsuphasawat, and Jeffrey Heer. 2016. Vega-Lite: A Grammar of Interactive Graphics. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (2016), 341–350. DOI: http://dx.doi.org/10.1109/TVCG.2016.2599030

[40] Arvind Satyanarayan, Ryan Russell, Jane Hoffswell, and Jeffrey Heer. 2015. Reactive Vega: A Streaming Dataflow Architecture for Declarative Interactive Visualization. *IEEE Transactions on Visualization and Computer Graphics* 22, 1 (2015), 659–668. DOI: http://dx.doi.org/10.1109/TVCG.2015.2467091

[41] Phoebe Sengers, Kirsten Boehner, Shay David, and Joseph'Jofish' Kaye. 2005. Reflective Design. In *Proceedings of the 4th Decennial Conference on Critical Computing: Between Sense and Sensibility*. ACM, 49–58. DOI: http://dx.doi.org/10.1145/1094562.1094569

[42] Tableau Software. 2019. Explain Data: AI-Driven Explanations for Your Data. https://www.tableau.com/products/new-features/explain-data. (2019).

[43] Arjun Srinivasan, Steven M Drucker, Alex Endert, and John Stasko. 2018. Augmenting Visualizations with Interactive Data Facts to Facilitate Interpretation and Communication. *IEEE Transactions on Visualization and Computer Graphics* 25, 1 (2018), 672–681. DOI: http://dx.doi.org/10.1109/TVCG.2018.2865145

[44] Sharifa Sultana and Syed Ishtiaque Ahmed. 2019. Witchcraft and HCI: Morality, Modernity, and Postcolonial Computing in Rural Bangladesh. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, Article Paper 356, 15 pages. DOI: http://dx.doi.org/10.1145/3290605.3300586

[45] Alice Thudt, Uta Hinrichs, and Sheelagh Carpendale. 2012. The Bohemian Bookshelf: Supporting Serendipitous Book Discoveries through Information Visualization. In *Proceedings of the 2012 CHI Conference on Human Factors in Computing Systems (CHI '12)*. Association for Computing Machinery, New York, NY, USA, 1461–1470. DOI: http://dx.doi.org/10.1145/2207676.2208607

[46] Manasi Vartak, Silu Huang, Tarique Siddiqui, Samuel Madden, and Aditya Parameswaran. 2017. Towards Visualization Recommendation Systems. *ACM SIGMOD Record* 45, 4 (2017), 34–39. DOI: http://dx.doi.org/10.1145/3092931.3092937

[47] Manasi Vartak, Samuel Madden, Aditya Parameswaran, and Neoklis Polyzotis. 2014. SeeDB: Automatically Generating Query Visualizations. *Proc. VLDB Endow.* 7, 13 (Aug. 2014), 1581–1584. DOI: http://dx.doi.org/10.14778/2733004.2733035

[48] Kaveh Waddell. 2019. The dangers of "AI washing". https://www.axios.com/ai-washing-hidden-people-00ab65c0-ea2a-4034-bd82-4b747567cba7.html. (2019).

[49] Arthur Edward Waite. 1999. *The Original Rider Waite Tarot Deck*. US Games Systems, Incorporated.

[50] Yun Wang, Zhida Sun, Haidong Zhang, Weiwei Cui, Ke Xu, Xiaojuan Ma, and Dongmei Zhang. 2019. DataShot: Automatic Generation of Fact Sheets from Tabular Data. *IEEE transactions on visualization and computer graphics* (2019). DOI: http://dx.doi.org/10.1109/TVCG.2019.2934398

[51] Leland Wilkinson. 2017. Visualizing Big Data Outliers Through Distributed Aggregation. *IEEE Transactions on Visualization and Computer Graphics* 24, 1 (2017), 256–266. DOI: http://dx.doi.org/10.1109/TVCG.2017.2744685

[52] Kanit Wongsuphasawat, Dominik Moritz, Anushka Anand, Jock Mackinlay, Bill Howe, and Jeffrey Heer. 2015. Voyager: Exploratory Analysis via Faceted Browsing of Visualization Recommendations. *IEEE Transactions on Visualization and Computer Graphics* 22, 1 (2015), 649–658. DOI: http://dx.doi.org/10.1109/TVCG.2015.2467191

[53] Emanuel Zgraggen, Zheguang Zhao, Robert Zeleznik, and Tim Kraska. 2018. Investigating the Effect of the Multiple Comparisons Problem in Visual Analysis. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 479. DOI: http://dx.doi.org/10.1145/3173574.3174053