APPENDIX

In this section we present the in-situ chart representations within the overview, NOAH s underlying system architecture, and statistical significance results from our user study.

EXPLAINING THE CHART REPRESENTATIONS

In this section, we explain the chart representation of all the formula categories discussed in Section 5, also displayed in Figure 4.

Summary. The result of a *summary* formula, *e.g.*, `AVERAGE`, is depicted using a horizontal bar. The bar represents the range of the data subset the bin spans, with the minimum and maximum values annotated within the chart. A vertical line is used to highlight where the result lies within the range.

Conditional. The result of a *conditional* formula, *e.g.*, `COUNTIF`, is depicted using a histogram. The histogram captures the distribution of the attribute on which the formula has been applied, *e.g.*, the *availability* attribute discussed in Section 2. Shading is used to de-emphasize data ranges that do not satisfy the condition. A vertical line is used to highlight where the result lies within the distribution.

Frequency. The result of a *frequency* formula, *e.g.*, `mode`, is also depicted using a histogram. The bin in the histogram that contains the result is rendered with "orange" color. A vertical line is used to highlight where the result lies within the distribution.

Spread. Finally, the result of a *spread* formula, *e.g.*, `mode`, is depicted using a histogram. A similar shading technique as the *conditional* formula is used to highlight the standard deviation. The mean is highlighted using a vertical line.

In the chart representation mode, each entry of the aggregate column contains one additional visual cue—a color bar with shades of green on the right of the chart (see Figure 3). The darker the color, the higher the value corresponding to that entry. Users can use the color intensity to compare results among different aggregate column entries.

IMPLEMENTATION AND ARCHITECTURE

In this section, we provide an overview of the infrastructure of NOAH. We integrate NOAH as a data exploration plugin within DATASPREAD [?], an open-source scalable web-based spreadsheet.

Underlying Data Structures

As explained in Section 5, the underlying data structure representing the overview is an in-memory equi-depth histogram. NOAH constructs the histogram on-demand based on the navigation attribute. In the beginning, only the highest granularity bins are constructed. As users perform ad-hoc interactions on the data, the interface is updated on the fly. For example, when a user zooms into a specific bin, NOAH again constructs an equi-depth histogram on the data corresponding to that bin on demand. To enable seamless integration of the overview with the spreadsheet data, we leverage the hierarchical positional indexes used by DATASPREAD [?] to access the spreadsheet data. The index is essentially an order statistics tree [?] built on the position (*e.g.*, row number) of the spreadsheet data. For any given navigation attribute, a new positional index is constructed first. NOAH then leverages the positional mapping to access the underlying data corresponding to the navigation attribute and constructs the histogram depicting the overview. Each bin in the histogram maintains positional information regarding its elements, *i.e.*, starting and ending index of each unique element in the bin (*e.g.*, cities in the Airbnb data). Therefore, NOAH can be integrated into any spreadsheet and requires only access to the positional mapping structure of that spreadsheet.

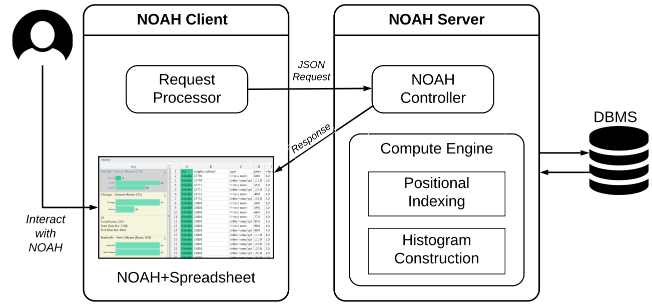

Fig. 8. NOAH System architecture.

System Architecture

We now explain the system architecture of NOAH. The NOAH client is a web-based front-end that captures user input and renders both the navigation plug-in and the spreadsheet based on the results returned by the back-end. The front-end is responsible for capturing user input and rendering components of the navigation interface, *i.e.*, the overview, aggregate column, and context bar. Given any interaction by the user on the front end, the *request processor* issues a request to back-end. The back end navigation controller receives the request from the front-end. After processing the request that corresponds to some front-end user interaction, the *request processor* sends a response to the front-end encoded in *json*. For requests involving the spreadsheet, *e.g.*, scrolling, the *request processor* leverages the *positional index* to access the spreadsheet data. For requests involving the navigation interface, *e.g.*, zoom in/out, the *request processor* leverages the *compute engine* to manage the equi-depth histogram on demand. The compute engine is also responsible for processing analytical operations. We leverage DATASPREADs built-in formula engine to support the analytical operations.

RESULTS

In this section, we evaluate the statistical significance of the task performance results and survey responses presented in Section 7.

Intra-participant differences.

Figure 9 depicts the intra-participant submission time differences between NOAH and Excel, across all the quiz tasks. For a given task and a participant, the corresponding circle denotes by what percentage (between 0.30% to 94.33%) a

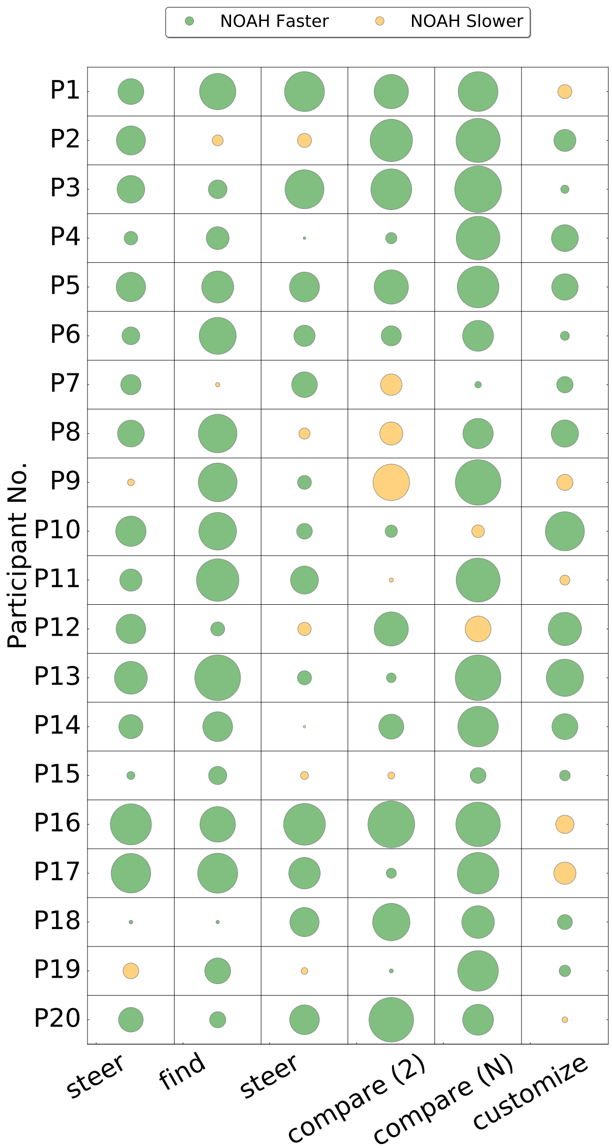

Fig. 9. Intra-participant submission time differences between NOAH and Excel across the quiz tasks.

participant on a given tool was faster than on the other tool, when submitting a response. The color of the circle denotes on which tool the participant's submission time was faster (green: NOAH faster, yellow: Excel faster). The larger the circle, the faster the corresponding tool is. For example, participant $P9$'s submission time was 91.78% faster with NOAH for the compare (N) task and 72.62% faster with Excel for the compare (2) task. Figure 9 confirms that majority of the submission times using NOAH were faster compared to Excel—nineteen out of the twenty participants completed at least four tasks in less time using NOAH compared to

For compare (2), customize, and the second steer task, fewer than a third of the participants' submission times were faster using Excel. For the second steer tasks, four out of the six participants that submitted answers quickly using Excel compared to NOAH used the *autosum* shortcut for the birdstrikes dataset. For the compare (2) tasks, a number

Excel. The submission time difference is even more apparent for the steer, find, and compare (N) tasks.
of participants ($N = 3$) that required more time to submit an answer using NOAH, first used the bin customization feature to view all the unique bins, contributing to a higher submission time compared to Excel. On the other hand, as explained earlier, the unfamiliarity with bin customization contributed to higher submission time using NOAH for six participants (yellow circles).

Statistical Significance: Task Performance

Since we conducted our user study on a relatively small population (20 participants), we further evaluated the statistical significance of the task performance results, *i.e.*, accuracy and completion time. To measure the significance of the task completion times, we ran *Mann-Whitney's U test* (as completion times did not follow a normal distribution). *For all of the tasks except the customize task, we found a significant effect of the tools, i.e., the response times for the tasks significantly differed by the choice of the tool* (see Table 3). We ran the *Fisher's exact test* that measures the statistical significance of categorical data to asses the difference in accuracies (0/1 accuracy) : *the percentage of accurate of submissions significantly differed by the choice of the tool only for the steer tasks*.

TABLE 3
Statistical significance of submission time and accuracy comparisons between NOAH and Excel. (*) indicates statistically significant.

Question	Category	Time (p value)	Accuracy (p value)
$Q1$	steer	**0.0007** (*)	**0.0033** (*)
$Q2$	find	$\mathbf{2.49 \times 10^{-5}}$ (*)	0.7475
$Q3$	steer	**0.0043** (*)	**0.0202** (*)
$Q4$	compare (2)	**0.0154** (*)	1
$Q5$	compare (N)	$\mathbf{5.83 \times 10^{-6}}$ (*)	0.48
$Q6$	customize	0.1207	0.0959

Statistical Significance: Subjective Ratings

We further conducted a statistical significance test—the *Wilcoxon Signed-rank* test—on the survey responses, with results shown in Table 4. For all the metrics, the ratings significantly differed by the choice of the tool, *i.e.*, NOAH or Excel. The distribution of the ratings for none of the criteria followed a normal distribution.

TABLE 4
Survey results. (*) indicates statistical significance.

Metric	NOAH	Excel	p value
Ease of Learning	$\mu = 5.75$, $\sigma = 1.02$	$\mu = 4.22$, $\sigma = 1.41$	$\mathbf{1.49 \times 10^{-7}}$ (*)
Speed of Use	$\mu = 6.03$, $\sigma = 0.99$	$\mu = 4.22$, $\sigma = 1.65$	$\mathbf{1.68 \times 10^{-7}}$ (*)
Ease of Use	$\mu = 5.88$, $\sigma = 0.90$	$\mu = 4.33$, $\sigma = 1.71$	$\mathbf{7.85 \times 10^{-6}}$ (*)
Confidence	$\mu = 5.50$, $\sigma = 1.79$	$\mu = 4.60$, $\sigma = 1.50$	**0.0096** (*)
Comprehensibility	$\mu = 5.60$, $\sigma = 1.27$	$\mu = 4.48$, $\sigma = 1.65$	**0.0006** (*)
Satisfaction	$\mu = 5.48$, $\sigma = 1.16$	$\mu = 4.52$, $\sigma = 1.49$	**0.0018** (*)