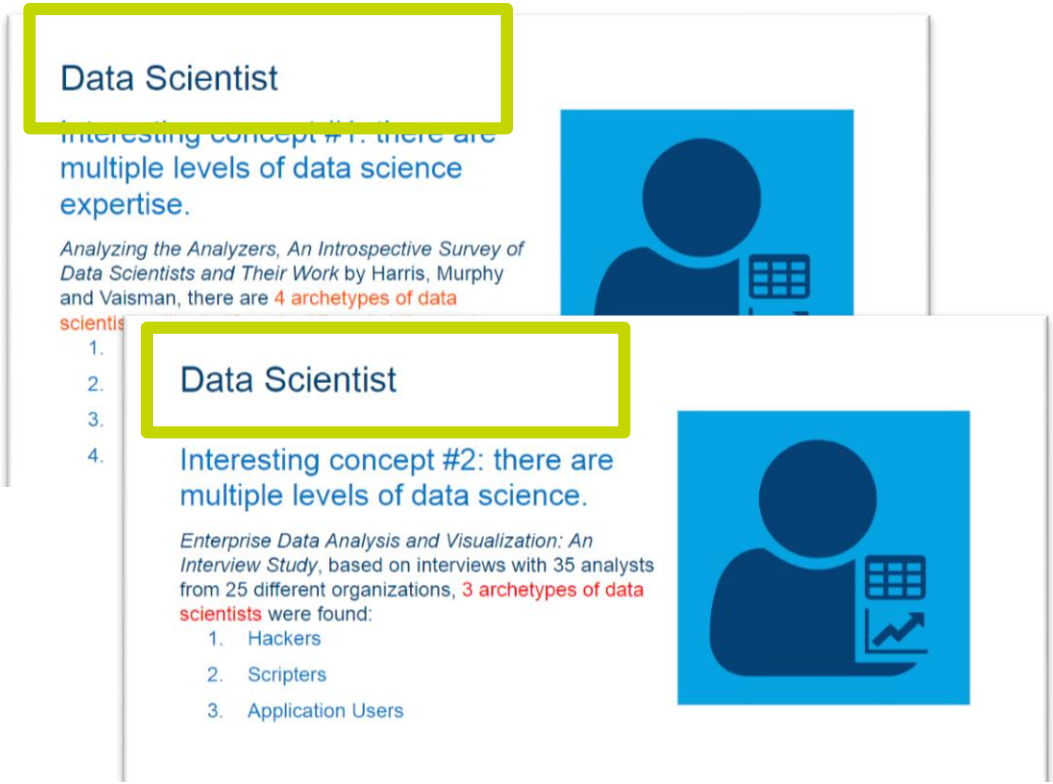
Defining Data Scientists

How do I define a role that is emerging?

Defining Data Scientists



Defining Data Scientists

Defining Data Scientists

EXPERIMENTATION SEGMENT

- Novice-Mid-level data scientists
- Have been told that ML/DL is the latest and greatest thing and they need to learn it
- Typically don't have a domain specific problem that could be solved with ML/DL
- Usually starts with one model ("I read that model x will solve problem y, I am going to try this.")
- Often start without data and uses a data set from online to learn the patterns of data science
- Might move to deploying a model for use but typically in a demo type situation, experimenting with the possibilities of ML/DL
- Coding Style - Hacker

Defining Data Scientists

SOLUTIONS SEGMENT

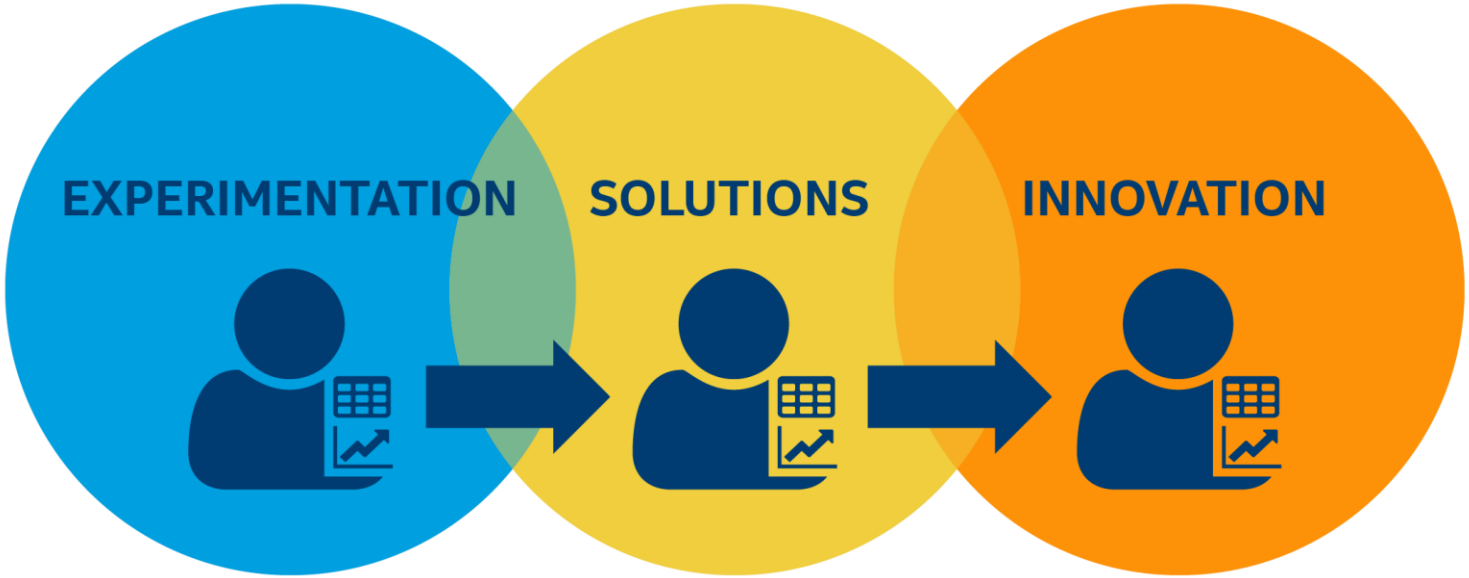
- Mid-senior level scientist
- Have experimented or used ML/DL in a real world application
- Have researched ML and DL and have an idea of what model to use to solve problems
- Have a domain specific problem that needs to be solved with ML/DL
- Knows that they can’t solve the problem without the correct data set
- Will often combine/adjust models to solve a specific problem
- Deploys models for use in the real world
- Coding Style - Hacker

Defining Data Scientists

INNOVATION SEGMENT

- Senior level scientist
- Data Science degrees Masters level or PhD level, etc.
- Have deep knowledge of various models to use for specific problems
- Have a domain specific problem that needs to be solved with ML/DL
- Is looking for the latest and greatest model and is well read on what is available
- Knows they can't solve the problem without the correct data set
- Deploys models for use in the real world
- Will combine/adjust models to solve problems
- Coding Style - Programmer

Defining Data Scientists



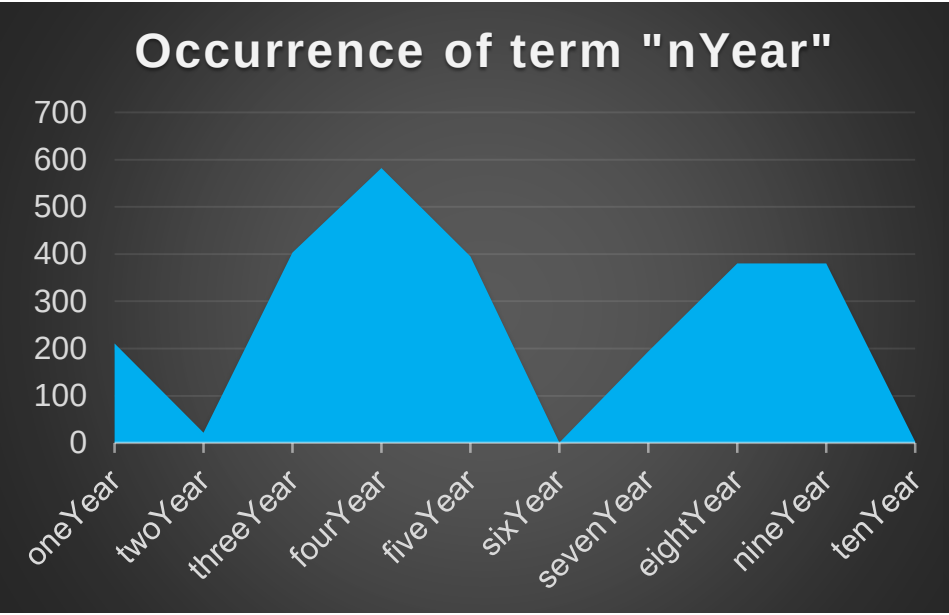
My own data scrape results

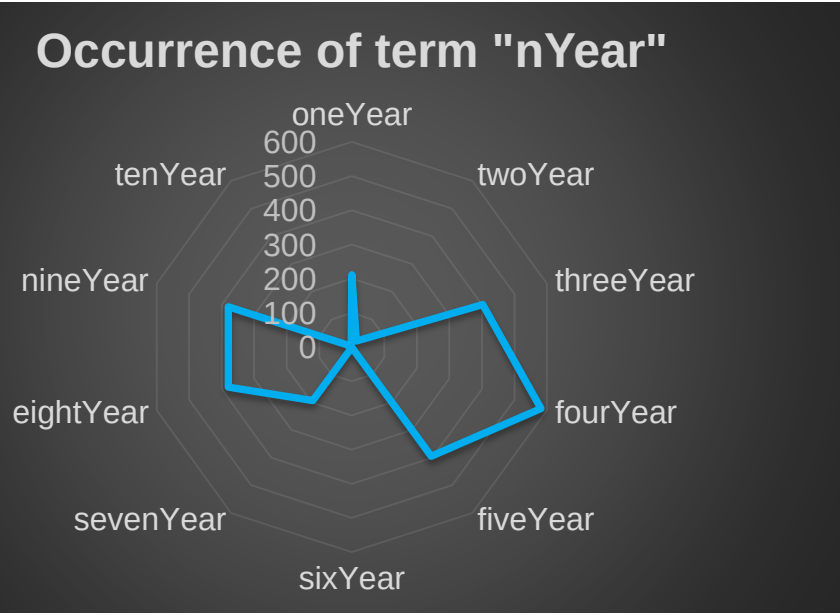
Methodology:

- Scraped Indeed.com only
- Scraped for “Data Scientist” job postings only
- Scraped for common data science tools
- Scraped for different levels of jobs (entry level - senior level)
- What employers expect, might give us some insights into the tools that different data scientist levels are using.

Data Scrape Results

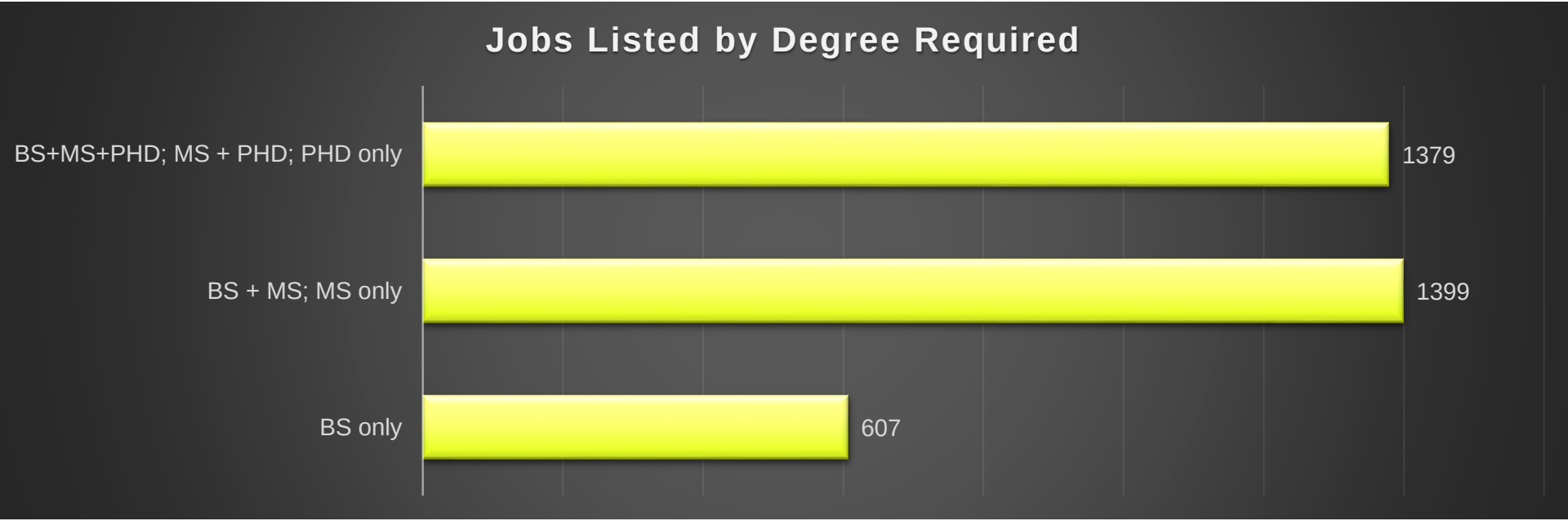
Scraped for terms “nYear” (oneYear, twoYear, threeYear, fourYear, etc.)





Data Scrape Results

Scraped for terms “BS, MS, PHD” (Bachelors, BS, Masters, Masters Degree, etc.)



Goal of Data Scrape

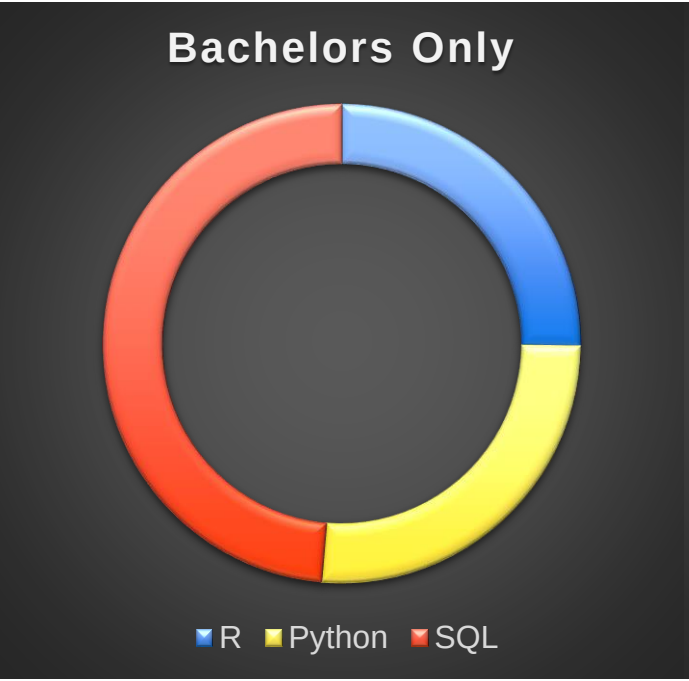
Disprove the theory that the higher the level of data science, the more likely it is they will use CLI based tools rather than UI based tools.

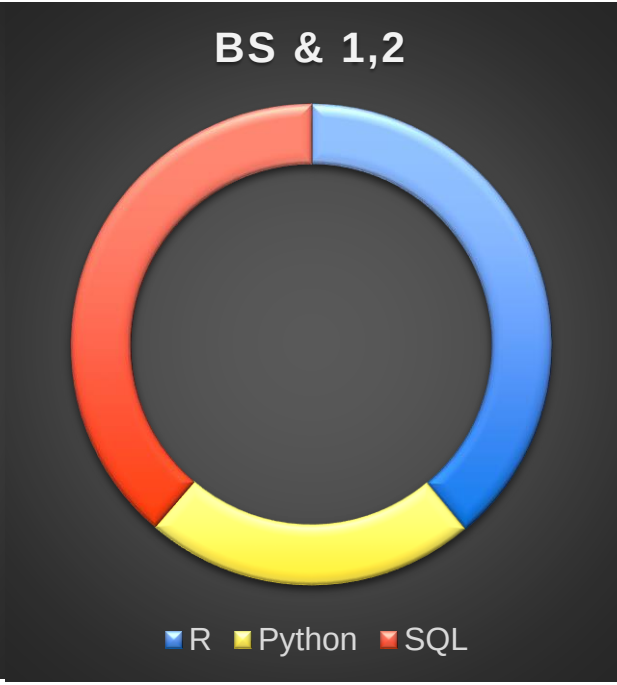
Scraped for tools in groups:

- Frameworks
- Languages
- UI based tools

How to calculate experience?

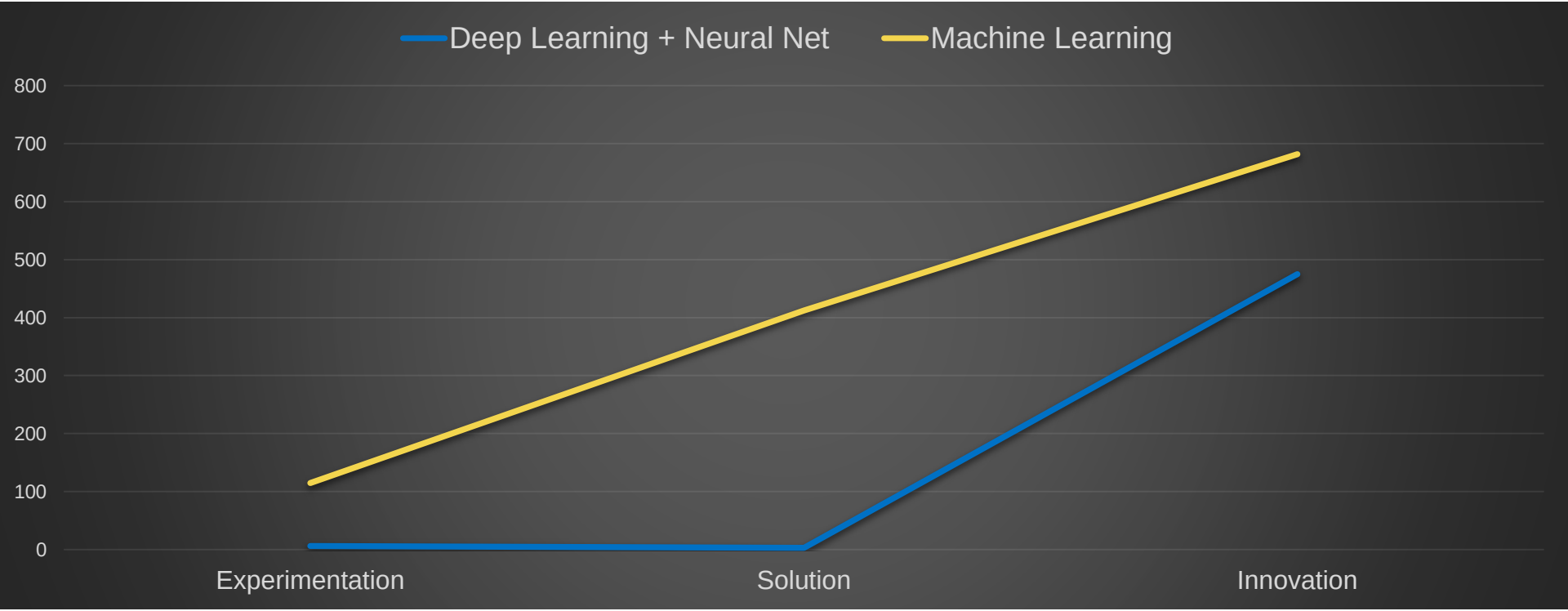
Average of degree totals with experience totals.





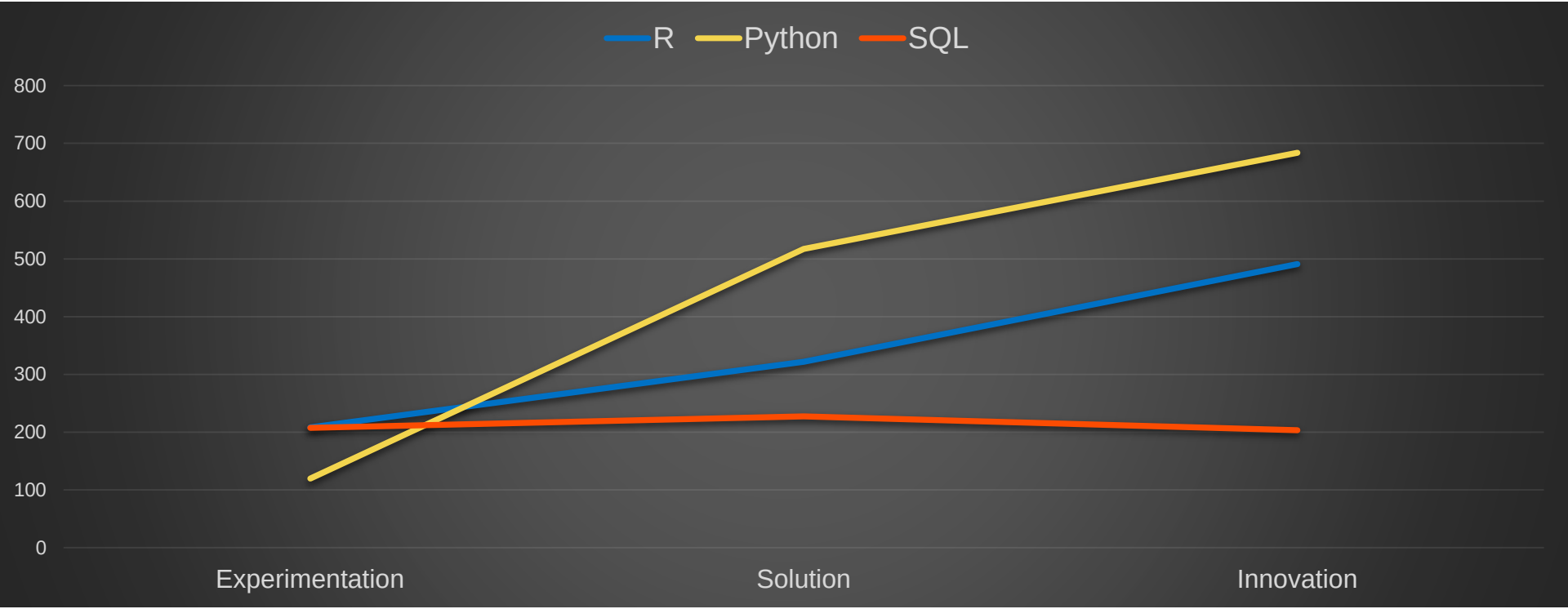
Data Scrape Results

Terminology by Data Science Level



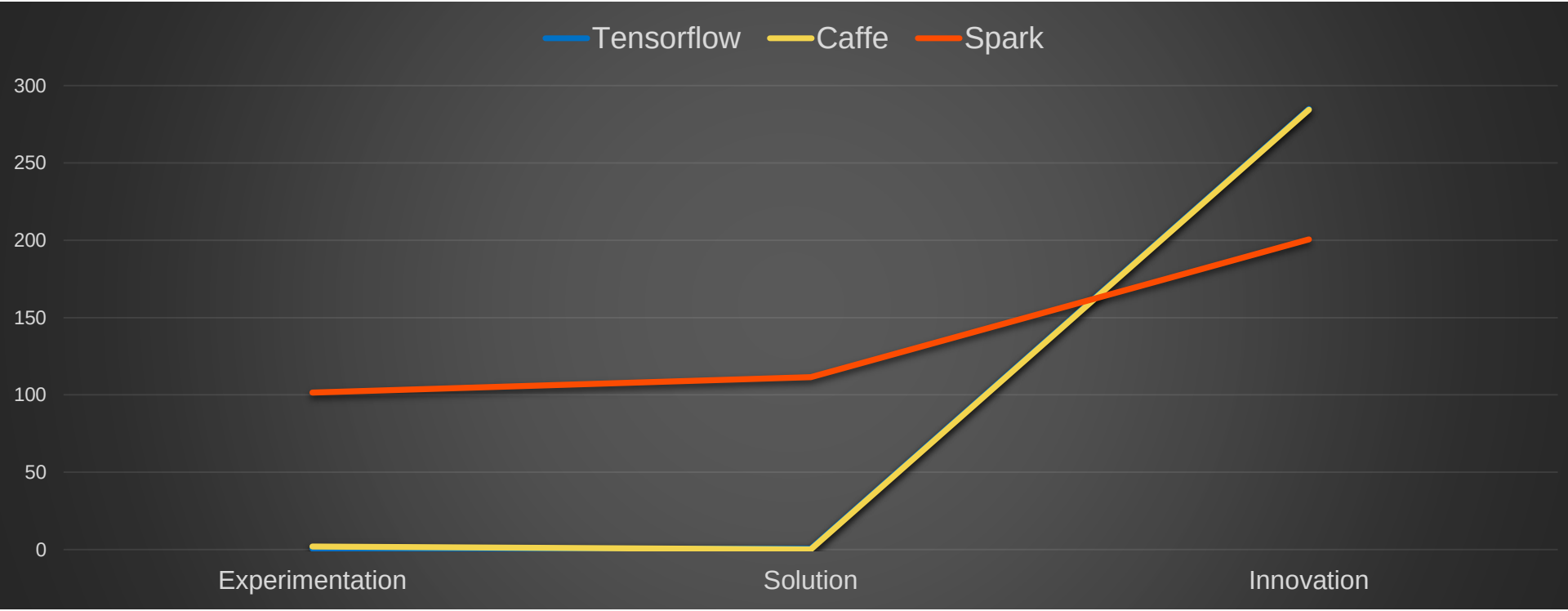
Data Scrape Results

Languages by Data Science Level



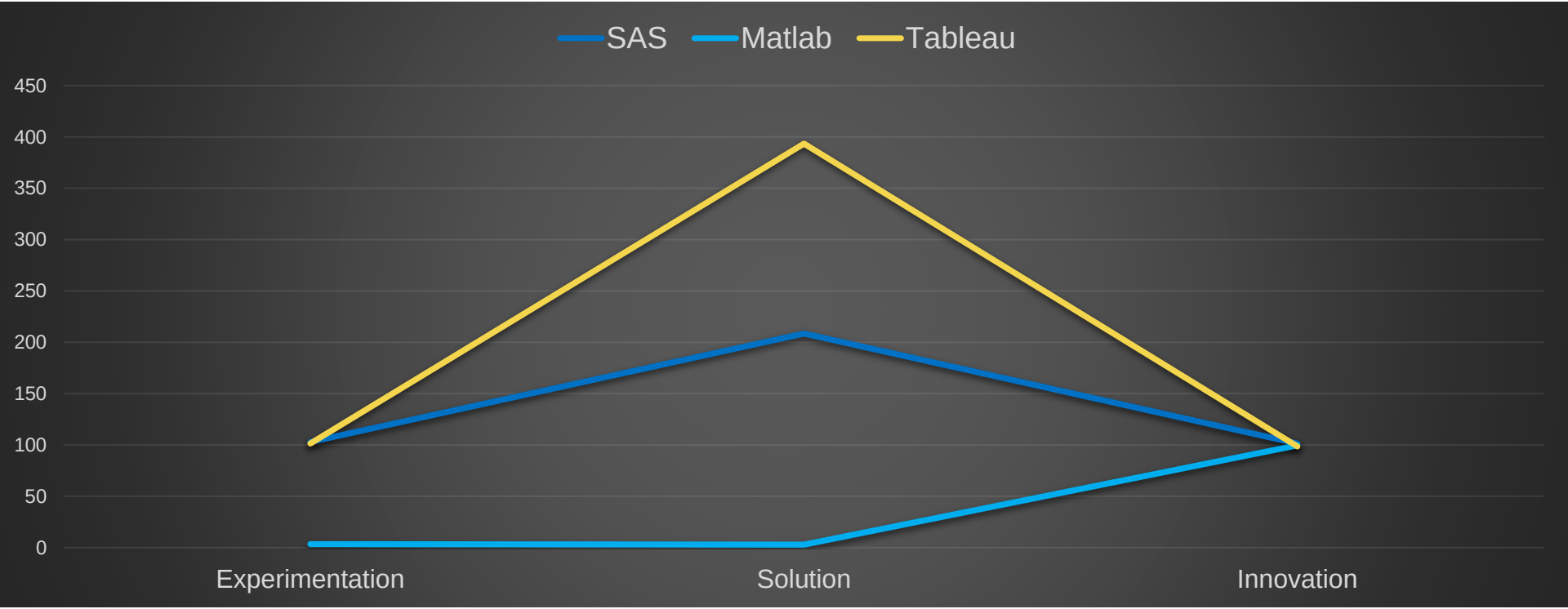
Data Scrape Results

Frameworks by Data Science Level



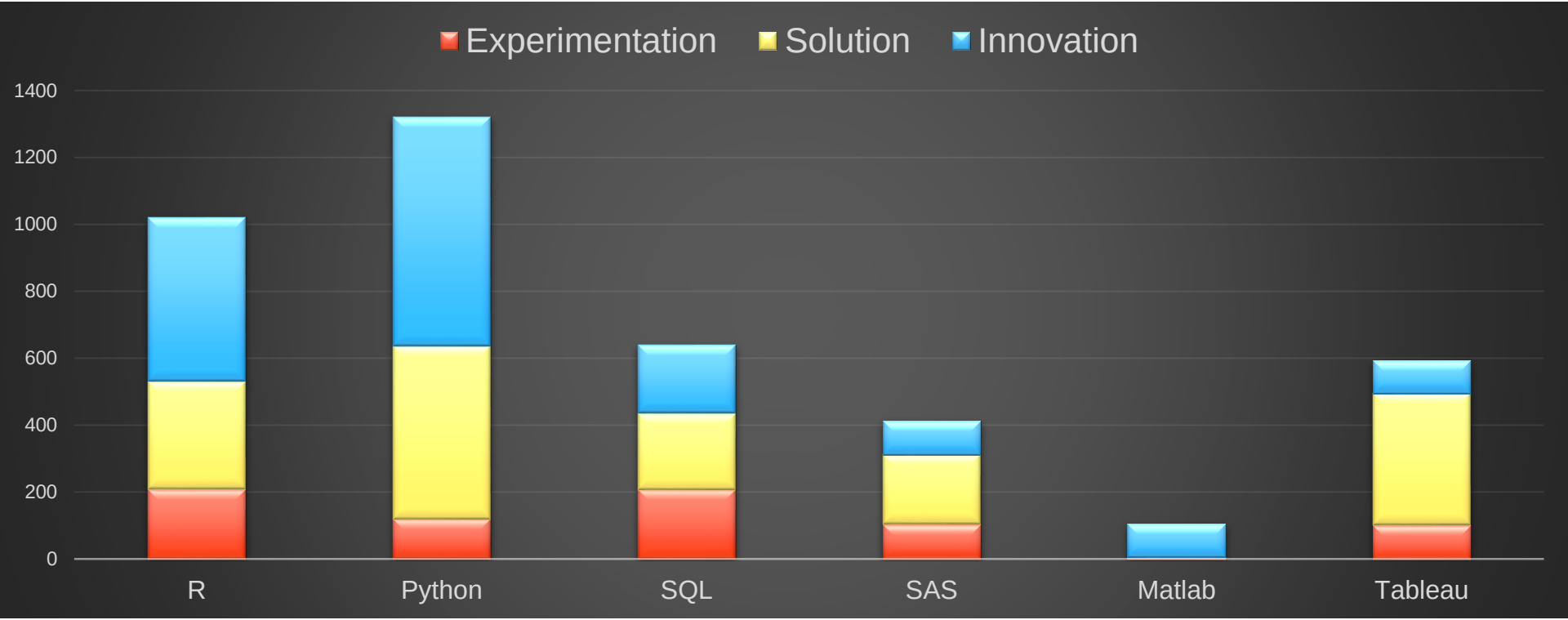
Data Scrape Results

Tools with UI by Data Science Level



Data Scrape Results

Tools with UI vs CLI by Data Science Level



Data Scrape Conclusions

1. CLI tools occur at a higher rate than UI based tools.
2. The solutions level uses UI based tools at a slightly higher rate than the experimenter level and innovator level.
3. Caffe and Tensorflow are neck an neck in popularity, and required less than in experimenter level and solutions level than in the innovator level. This could be due to the innovators affinity for finding the latest and greatest product.

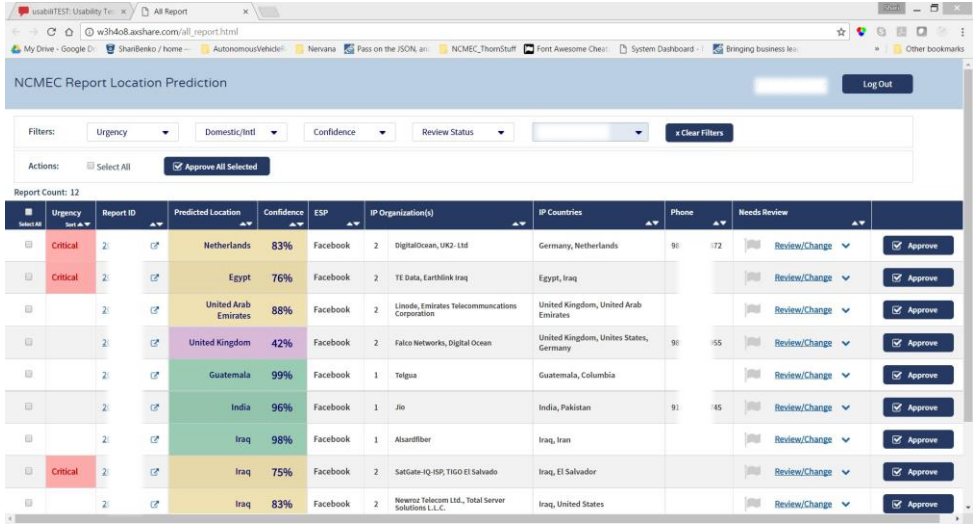
NCMEC PREDICTION LOCATOR USABILITY TESTING RESULTS

Prototype v8 – 6/25/2017

Presented by: Shari Benko, UX Designer

TEST STRUCTURE

- Users shared their screens and were given a URL to v8 of the working prototype.



- **2 questions** & **8 tasks** were given to **5 participants** (**2 inexperienced** NCMEC analysts and **3 experienced**)
- SUS survey given to each participant at the end of the tests

NCMEC
Location Prediction
Prototype v8
Usability Testing 6/14 - 6/20/2017

QUESTION & TASK ANALYSIS

Completion rates and times

QUESTION #1

What is your first impression of this queue? What is its purpose?

- **Reasoning:** This is to gain qualitative feedback on first entrance into queue.
- **Prediction:** There is no predicted value for qualitative feedback.

Results:

- Subject 1 (inexperienced analyst) was confused and overwhelmed at first
- Subjects 2-5 jumped to answering question #2

Conclusions:

- The queue could seem overwhelming to inexperienced analysts, but wasn't troubling to users who were comfortable with their current workflow

QUESTION #2

Which reports would you start with first and why?

- **Reasoning:** To discover which items a user will tackle first.
- **Prediction:** Each analyst will have their own way of choosing which reports to start with first.

Results:

- 3 out of 5 started with either high or low confidence reports
- 2 out of 5 started with Critical urgency, then moved to confidence levels

Conclusions:

- Confidence was clear to the end users as the place to start urgency was a close 2nd

TASK #1

How would you forward a report to the predicted jurisdiction?

- **Reasoning:** To assess how intuitive the “Approve” button is to find and use.
- **Prediction:** User should indicate that “Approve” is the correct button to forward a report.

Results:

- **Task Completion Rate:** 80%
- **Task Average Completion Time:** 5.25 seconds (KLM = X)

Conclusions:

- The approve button was clear to most users as the way to forward a report from this queue

TASK #2

Which reports do you feel that you need to review before forwarding to the predicted jurisdiction?

- **Reasoning:** To assess whether or not the confidence levels are what indicate to the user that the report needs to be reviewed, and to assess the content presented in the unexpanded view.
- **Prediction:** The user should indicate reports that need to be reviewed based on the confidence level, or content presented in the unexpanded view.

Results:

- Users were presented with dummy information which made the 2nd prediction impossible to evaluate quantitatively
- 2 of the users concluded that the low confidence was the indication

Conclusions:

- This task is inconclusive and no conclusions can be made

TASK #3

How would you save a report for review at a later time?

- **Reasoning:** To find out how quickly a user understands that they can flag a report for review and it drops the report to the bottom of the list.
- **Prediction:** The user should indicate or click on the flag, to flag a report for review at a later time.

Results:

- **Task Completion Rate:** 20%
- **Task Average Completion Time:** 33 seconds

Conclusions:

- The review flag button is not obvious enough to users

TASK #4

Which reports do you think have an accurate enough prediction to forward without review?

- **Reasoning:** To assess how intuitive the wording “confidence” is to an analyst, and how much they trust the system’s predictions.
- **Prediction:** User should indicate that the high confidence reports are accurate enough to forward.

Results:

- 2 out of 4 users indicated that high confidence
- 3 users indicated they would trust the system over time

Conclusions:

- While completion rate and times couldn’t be calculated due to dummy information, the users indicated they would trust the system over time

TASK #5

How would you forward multiple reports at the same time, that don't need review?

- **Reasoning:** To assess how quickly it is for a user to filter and approve multiple reports at once.
- **Prediction:** User will filter for items that indicate an accurate prediction, then check select all either from the action menu or from the table header, then select the button for "Approval All Selected.

Results:

- **Task Completion Rate:** 100%
- **Task Average Completion Time:** 15 seconds
- **1 out of 5 users used filters to complete the task**

Conclusions:

- All users completed the approve all task using different methods, but were successful indicating the quick actions is intuitive to new users
- Filters were not an obvious interactions to users

TASK #6

How would you go about reviewing reports that do need review before forwarding?

- **Reasoning:** To assess whether or not the expand/close interactions are easy to find.
- **Prediction:** The user should click the "Review/Change" down arrows to expand the report to start a review.

Results:

- **Task Completion Rate:** 40%
- **Task Average Completion Time:** 5 seconds
- **Subject 1 eventually clicked review after being prompted a 2nd time**
- **Subject 2 attempted to use unexpanded info to review**
- **Subject 5 clicked through to external report to review**

Conclusions:

- Since 3 out of 5 eventually reached the expanded view – but 2 used info on screen or external, this functionality while somewhat intuitive needs training

TASK #7

How would you change the jurisdiction of a report?

- **Reasoning:** To discover how intuitive it is to change the location of a report.
- **Prediction:** The user will click the “Review/Change” expander, select the drop down box for change location, select a new location and interact with the modal window.

Results:

- **Task Completion Rate:** 100%
- **Task Average Completion Time:** 10.4 seconds
- **All subjects understood the purpose of the feedback box**

Conclusions:

- Changing a location of a report is easy for users to find
- The feedback modal window is easy for users to understand

TASK #8

Your co-worker Casey Blake is out sick today and you have been given her reports to handle. How would you start working on his assigned reports?

- **Reasoning:** To assess how easy it is for a user to work on a co-workers assigned reports.
- **Prediction:** User should select the filter for Casey Blake

Results:

- **Task Completion Rate:** 80%
- **Task Average Completion Time:** 6.25 seconds
- **Subject 5 looked for an export reports feature**

Conclusions:

- Most users had no trouble reaching a co-workers reports

OVERALL COMPLETION RATE AND TIME

Average completion rate: 70%

Average completion time: 10.5 seconds

What rate is a good rate? An analysis of almost 1200 usability tasks shows that the average task-completion rate is 78%. https://measuringu.com/sus/

SYSTEM USABILITY SCALE SURVEY

Each test subject was given a link to a SUS survey to complete after the user testing session. The following cases were given to each subject along with a scale ranking and each users answers are then calculated into an overall score:

- I think that I would like to use NCMEC Prediction Locator frequently.
- I found NCMEC Prediction Locator unnecessarily complex.
- I thought NCMEC Prediction Locator was easy to use.
- I think that I would need the support of a technical person to be able to use NCMEC Prediction Locator.
- I found the various functions in NCMEC Prediction Locator were well integrated.
- I thought there was too much inconsistency in NCMEC Prediction Locator.
- I would imagine that most people would learn to use NCMEC Prediction Locator very quickly.
- I found NCMEC Prediction Locator very cumbersome (awkward) to use.
- I felt very confident using NCMEC Prediction Locator.
- I needed to learn a lot of things before I could get going with NCMEC Prediction Locator.

For more information on SUS scores, visit: https://measuringu.com/sus/

SYSTEM USABILITY SCALE SURVEY RESULTS

SUS score: **81.3** (percentile: 91%, which corresponds to the following letter grade: B-)

Learnability: **84.4**

Usability: **80.5**

Standard deviation: **11.3**

Cronbach's alplha: **0.84** (above 0.7 is acceptable)

What score is a good score? A SUS score above a 68 would be considered above average and anything below 68 is below average. https://measuringu.com/sus/

CONCLUSIONS AND RECOMMENDATIONS

What do we do based on this test?

CONCLUSIONS & RECOMMENDATIONS

The queue could seem overwhelming to inexperienced analysts, but wasn't troubling to users who were comfortable with their current workflow

Recommendations: No changes needed

Confidence was clear to the end users as the place to start urgency was a close 2nd

Recommendations: No changes needed

The approve button was clear to most users as the way to forward a report from this queue

Recommendations: No changes needed

CONCLUSIONS & RECOMMENDATIONS

The review flag button is not obvious enough to users

Recommendations: rethink the necessity of this interaction

Users indicated they would trust the system over time

Recommendations: no changes needed

All users completed the approve all task using different methods, but were successful indicating the quick actions is intuitive to new users

Recommendations: no changes needed

A user was misled by the check icon on the Approve button

Recommendations: remove the check icon from the approve buttons

CONCLUSIONS & RECOMMENDATIONS

Filters were not an obvious interactions to users

- **Recommendations:** No changes needed, this interaction will be learned through use of the system and training on the system

Change/Review expand interaction required more time to learn

- **Recommendations:** No changes needed, this interaction will be learned through use of the system and training on the system

Changing a location of a report is easy for users to find

- **Recommendations:** No changes needed

CONCLUSIONS & RECOMMENDATIONS

The feedback modal window is easy for users to understand

- **Recommendations:** No changes needed

Most users had no trouble reaching a co-workers reports

- **Recommendations:** No changes needed

FINAL RECOMMENDATIONS

- Rethink the necessity of the save and review later flag interaction

  Discover whether or not this is an interaction that only a small fraction of people will use, or include a “Save for Later” interaction in the actions menu removing the flag option.

- Remove the check icon from the approve buttons

ESTIMATED COMPLETION RATE AFTER CHANGES

This estimate is calculated on the following assumptions:

- The completion rate for the “save for review later” feature is at least 80% once changes have been made.
- With a small amount of training which will take place due to the extensive training conducted by NCMEC for analysts, using the filters and using the expanded view for review should both increase to at least 80%.

Estimated completion rate: 86.6%

What rate is a good rate? An analysis of almost 1200 usability tasks shows that the average task-completion rate is 78%. https://measuringu.com/sus/

intel®

experience
what's inside™