Good Day Danielle,

I want to report on my experience with testing the existing and potential product data that we worked on in a previous assignment in R and Rstudio versus previously working in WEKA. As you might have guessed I Initially I did have some trouble trying to install R and RStudio in that the software did not appear to work at all, but after some troubleshooting of trying different download sources online and realizing that the R software needed to be installed before RStudio things eventually came out okay. Thankfully I do have some previous experience with installing a variety of different software packages, so these problems were not new to me. Truth be told it really did just came down to trial and error to find the right way of getting software installed correctly.

What I found invaluable with starting in RStudio was in going through the programming tutorial in chapter 14 of *Predictive Analytics for Dummies*. From my experience reading a programming book and experimenting with the code or going through the walkthroughs the writer presents is a far superior learning experience than just reading the book alone. How would a new programmer expect to understand what code compiling failures are like? What would they do if they did not know why a line of code failed to produce the results they expected? And my answer is that you must get hands on experience with programming. No matter how good of an instructional book someone has will not substitute for knowing how to troubleshoot issues that come up or to deal with the frustration of disappointing results. So would I recommend going through the tutorial? Yes, without hesitation.

Now the big money question is after going through the trouble of introducing myself to R and RStudio and experimenting with the datasets: Did I learn anything of value for the business? The answer is yes and no. Yes I did learn some interesting things about the existing and potential product data that made me think about it in a different way. Namely that the various star reviews have an interesting relationship to the volume attribute. Which lead me to create several linear models and conclude that the five potential products I recommended through WEKA may not all be the best choices based on their net profit! As you can see in the graphs below there are the items highlighted in green and yellow. The items in yellow (175, 180, 196, & 199) do not match with each other between the potential top five products produced by WEKA and R. It is also interesting to note that the items that do match (186, 176, & 187) have dramatically different profits and that the product types have changed from including smart phones and game consoles.

|  |  |  |  |
| --- | --- | --- | --- |
| **Potential Top 5 Profitable Products - WEKA** | | | |
| **Profits of Each Potential New Product** | **Product #** | **Brand Name** | **Product Type** |
| $37,236.80 | 186 | Apple | Tablet |
| $37,888.48 | 199 | Sony | Game Console |
| $40,459.76 | 176 | Razer | Laptop |
| $48,576.00 | 196 | Motorola | Smartphone |
| $280,032.80 | 187 | Amazon | Tablet |

|  |  |  |  |
| --- | --- | --- | --- |
| **Potential Top 5 Profitable Products - R** | | | |
| **Profits of Each Potential New Product** | **Product #** | **Brand Name** | **Product Type** |
| $43,701.66 | 175 | Toshiba | Laptop |
| $56,389.35 | 180 | Acer | Netbook |
| $74,162.17 | 186 | Apple | Tablet |
| $151,660.23 | 176 | Razer | Laptop |
| $326,938.16 | 187 | Amazon | Tablet |

This makes me question if more analysis of the data is needed, which is not a bad thing. After all we are data analysts and trying to determine the best answer by viewing the data from different perspectives is what we do. And I should mention that that does not mean that the R results are wrong. I would just want to verify my results through further testing before presenting them to upper management.

I believe you had a question about my experiences getting into R. It is hard to describe what I have gained in knowledge from this experience I think in part because of how new things are for me in using R and Rstudio as compared to WEKA. For me I am a more visual leaning person, so at first I was apprehensive at trying my hand at another programming language, which to me in general comes down to just typing a lot of characters and symbols. I really like working in WEKA if for no other reason than I can just click through menus that allow me to go pretty quickly in changing features of classifiers and things of that nature while Rstudio at this point in my training takes a lot of time and effort to type something into existence.

From the start of this assignment I felt overwhelmed and unsure where to begin. In short, my work strategy was to take some time to mark down the various tasks I needed to do and work through them one by one until I got to the end. Yet through the whole experience I found I could not just mimic what was done previously in WEKA and have things work out fine like they tend to do for me. Instead I had to realize that I was not going to be able to work at this from the point of view of coming from WEKA and just implementing it directly into R. No, I had to think of the task from the R environment perspective to get results. For example little things like undoing choices in WEKA were gone, replaced with rerunning lines of code, requiring me to keep a notepad document of all of my code so that I did not have to recreate it from scratch. Even my volume results were different with some numbers being negative, which lead me to do research into why that must be. All told this project was not an easy task to accomplish, but I think I am better for it. While I would prefer to work in WEKA as I am more familiar and comfortable with it, I can see how R has its benefits.

If someone were starting with R there is some advice I would give based on my experience:

* **Give yourself time to learn the material**. For anyone new to programming it is important for them to understand that it is something that does not come easily. From this experience I learned that programming is a craft, something that has to be learned through trial and error in order to hone it into something useful and meaningful. Maybe because the code you are trying to run does not work or you cannot find an example right away that works with what you are trying to create. You have to give yourself plenty of time at the outset to make up for lack of knowledge and experience.
* **Find resources online or in print to help with discovering answers or examples of what you are working on**. I cannot count how many times and for that matter how many hours I spent scouring YouTube or using Google to search for more information on what I was working on from time to time. Use these resources as a starting point and then find specific sites that go that extra step in helping you achieve your goals. For me I often found help on sites like Qoura.com and Stackexchange.com.
* **Find a community for R programmers that can help answer your questions**. One thing that really stood out to me when working on this project was how isolated I felt. So I would recommend finding people online who are more experienced with R programming and reach out to them for answers and tips. And just knowing that other people faced the challenges I had was a good comfort to have.
* **Have patience and faith**. This tip ties into my advice for new programmers to give themselves more time to work than they may expect when starting with R. That there will come a time when something will just not make sense. When you have tried everything you can think and nothing seems to work. My best advice is to stop and step away from the computer for even just a few moments and allow your nerves to settle down because the fact of the matter is that R or any programming language is a foreign language to everyone starting out. And just like languages spoken in other countries they are difficult to learn because the thought process behind them is unique to that location. So, again, stop and take some time away to rest your mind and allow yourself to mentally reset and try again and believe you can do it.

Well Danielle, I hope this impromptu report was informative. As always if you have any questions let me know.

Regards

Edward Tanzer