RESOURCES AND NOTES FROM DIGITAL PORTFOLIO PRESENTATION

Use this Take Home Sheet from the Portfolio & Network Building (PNB) Group to explore the nuances of building your professional data science portfolio.

Follow along using the Slides linked here.

Slide 1 - Introduction - 1 minute

Slide 2 - Code of Conduct - 1 minute

Slide 3 - Introduction to the PNB - 2 minutes
- What is the PNB? The Portfolio and Network Building Group was created by the NEBDHub after several members of the CIC Student Working Group requested additional opportunities to meet other students and build their professional network. Some students wanted to learn about various data science projects available to them through the NEBDHub as well as methods for showcasing professional certificates in a digital portfolio.
  - Learn more on the PNB Website: https://nebigdatahub.org/pnb-group/
- The PNB Group represents the NEBDHub's reply to those requests. These workshops can help students build their professional data science portfolios, get ready for job interviews, polish their professional communications styles, and build their professional portfolio.
- The PNB Group hosts five sessions per year, each focusing on a different aspect of professional development, which you can leverage locally. The first session is for peer networking where attendees have the opportunity to meet one another and identify potential data science and professional collaboration opportunities in various domain areas from public health to environmental science. In the second session, we discuss job interviews and review resumes. In the third session, we review professional communication styles and best practices for presentations and cover letters. Today, we'll focus on building digital portfolios and the certificates students can earn from working on NEBDHub data science projects with real world data. Finally, we'll discuss Leadership Styles and how you can confidently lead teams in the workplace. The NEBDHub

encourages students to participate in all sessions that may be helpful to their professional development. This cycle repeats each year.
- We encourage you to host your own session using a downloadable toolkit available from the PNB website.
● Between meetings, you may join the PNB LinkedIn group where students share resources and career opportunities with one another.
- We also use the LinkedIn channel to host our mentorship conversations. Each week, PNB Leaders share a new prompt that poses a new question to the community, such as: "How can I differentiate myself in a tough job market?" or "How do I effectively discuss my research in an interview?". Our team of mentors who are professionals in academia and industry will weigh in, sharing their experiences.
- Definitely join the LinkedIn Group if you haven't already done so!

Learn more on the PNB Website: https://nebigdatahub.org/pnb-group/
Join the PNB LinkedIn: https://www.linkedin.com/groups/13009480/

Slide 4 - Today's Agenda - 1 minute
● But let's get going with today's PNB Session on Digital Portfolios. There's three parts to this workshop.
● We'll begin with a review of what Portfolios are, including why you need one as a data scientist and what hiring managers hope to see in your portfolio. Then, we'll work as a group to identify some of the projects you might want to put into your Portfolio.
● Second, we'll discuss the formatting of your portfolio and what platforms you can use to publish your projects.
● Next, we'll talk about what projects you can begin on for your portfolio. Emily will share her insights from the National Student Data Corps.
● Finally, we'll do a group Q&A and wrap up for the day. Feel free to drop your LinkedIn profiles in the chat to connect with one another throughout the presentation.

Slide 5 - What is a Data Science Portfolio? - 1 minute
● We will share a few slides here, but please drop your questions in the chat as we go - we'll do our best to answer them.
● To begin: What is a data science portfolio? A portfolio is an essential part of your job search toolkit.
● A portfolio is a complement to your resume, your LinkedIn, and your cover letter - it acts like a practical demo of your data science skills. A portfolio should showcase your project management experience as well as the range of data science skills you've learned either at University or from work or internships.

Slide 6 - Why do I need a Data Science Portfolio - 2 minutes
● Why do you need a Data Science Portfolio?

- A data science portfolio is becoming increasingly important in the data science job hunt. In Session B, back in the summer, we talked about technical interviews. Technical interviews are a part of the interview process where you might be asked to do either: 1) a live coding session or 2) a take-home assignment.
- Now, we can assume that hiring managers themselves are busy and low on time. A good data science portfolio can help them decide if they want to set aside part of their schedule to host a technical interview with you or not. For small companies, a solid portfolio might be enough for them to waive the technical interview altogether. Basically, having a good data science portfolio can be an advantage for you.
- The other personal benefit is that data science portfolios can help you monitor your professional growth over time. Use your portfolio to track your accomplishments!

Slide 7 - What goes in a Data Science Portfolio? - 3 minutes
- What goes in your portfolio?
- In short, you want to share any data science projects you've developed, particularly any projects that analyze real-world datasets. These can include projects you worked on in class or that we developed independently.
- If you are a beginner and are looking for data science projects to work on, we'll spend some time at the end of today's meeting talking about some DSPs you can work on through our National Student Data Corps program. So hang on for more info there!
- If you are a more advanced learner, you may want to explore datasets on your own and come up with your own research questions and projects.
  - Learn more about the data science projects available through the NEBDHub: https://bit.ly/nsdc-dsp
- One good way to develop your portfolio is to write down a list of the skills you'll need to qualify for your dream job. Find a job description that is aspirational but something you could reasonably qualify for in the next five years. Review their required skills and compare that list with your resume.
- If they want someone who is an expert in SQL, do a project on SQL. If they want someone who knows a lot about international supply chains, find a dataset on global commerce and start exploring. We encourage you to think about the technical skills and industry knowledge you'll need to be competitive in the job search and go from there.

Slide 8 - What do Hiring Managers want to see in a Data Science Portfolio? - 4 minutes
- But let's back up and talk about what hiring managers are looking for when they review your portfolio. Broadly speaking, hiring managers want to know four things:
  - Can you analyze data?
    - Do you have the technical skills to work with complex data? Do you also have the ability to properly analyze the data and see relationships between data points?
    - To answer this question, we want to include projects that use real world datasets and showcase a wide range of technical and critical thinking skills.

- ○ Can you write code?
  - ■ Your boss will want to know that you can not only analyze data, you can collaborate. Your insights are only valuable if others can easily double check your work and confirm your findings.
  - ■ To answer this question, we want to make sure our projects include the snippets of code used to, for example, extract data from your database. Think of this as show and tell - we want to practice writing neat and efficient code for our portfolio.
- ○ Can you write well?
  - ■ To be good at your job, no matter what job it is, you're going to have to be a good communicator. Even if your ideal job is highly technical, you have to be able to share your ideas with your immediate supervisor and coworkers. You have to be able to clearly structure and organize your thoughts. Good writing skills will help you do that.
  - ■ If you can learn to write well, you will stick out in the candidate pool. Learn to be a compelling communicator, and you'll be more successful at work.
- ○ Can you communicate your insights to a non-technical audience?
  - ■ Finally, we need to prove to a hiring manager that we can not only communicate, but we can communicate with non-technical audiences. In reality, when you're at work, you will be collaborating with people outside your department. You'll need to be able to explain your insights to the marketing department, the legal team, or operations techs.
  - ■ Your portfolio is a great place to showcase your ability to communicate effectively and at a very high level. Write about your work so that someone could understand your insights after reading for a minute or less. Use a mix of visualizations to illustrate your points. When possible, make recommendations that might support a business' needs (e.g., say when a finding might help you lower costs at a given company)
- ○ Before we move on, does anyone have any questions? We will share an additional resource on what hiring managers are looking for in a portfolio via the chat. We encourage you to take a look at that when you have the time.
  - ■ Additional resource on what a hiring manager is looking for in your portfolio here: https://youtu.be/7_VIjcFKHZo?feature=shared&t=163

Slide 9 - Data Science Portfolio Wish List - 3 minutes
- So if your portfolio can answer those four questions, you're off to a great start. There are a few other small things that will help you stand out.
- First, be professional but also let your personality shine through in your portfolio. If you're interested in sustainability, work on projects related to climate change. If you love creating complex data visualizations, work on projects that let you explore data journalism and use big splashes of color.

- Next, make sure your portfolio has a similar tone to your resume and cover letter. If your other application materials are formal and well constructed, your portfolio should be too. Think of all of your application materials as a package set!
- Aim for quality over quantity. It's better to have 5 excellent projects than 50 mediocre projects. A well-written and thoughtful project is better than a thrown together and confusingly formatted project.
- If you decide to put group projects in your portfolio (this is very common), be really clear about what you did vs. what your colleagues did. Don't make the hiring manager guess about your skills.
- Finally, make sure your Portfolio is accessible to everyone. Most websites now follow U.S. government guidelines about color contrast, alt text, and large text size. The NEBDHub, the parent organization of the Portfolio & Network Building (PNB) Group has a guide on how their own website accessibility plan was developed that we will share in the chat. This should help you get started.
    - View our accessibility policy here: https://nebigdatahub.org/accessibility/

Slide 10 - Job Description Exercise - 3 minutes
- So let's take a break from my slideshow and work on a group exercise. We will share in the chat a link to a sample job description for a Data Analyst role at Columbia University.
- We'll take a few minutes to read through this description silently. At [time], we'll come back together and discuss the following:
    - **Identify 3 technical skills** that you would need to grow to be competitive for that role (e.g. Python or Tableau).
    - The job's **domain area** is higher education. Consider what kinds of questions leaders in education are facing right now. What kinds of data can be used to help them answer those questions? **Identify 1 type of data or source of data that you could use to develop a research project in higher education.** This would help you prepare for a potential interview.
        - For example, would data students' declared majors be helpful to explore in preparation for this interview? (No, probably not).
    - Finally, what clues does this job description provide about the communication skills you'll need to be successful in this role? **Identify 1 communication-related responsibility you'd need to qualify for this job. Think about how you could showcase that skill in a portfolio project.** For example, if you need to create quarterly reports, you'd want to showcase some experience presenting data analysis.
- Sample job description here: https://nebigdatahub.org/wp-content/uploads/2025/01/Data-Analyst-Morningside-New-York-United-States.pdf

Break for **3 minutes**

NEBDHub's Portfolio & Network Building Group
Session D: Digital Portfolios
https://nebigdatahub.org/pnb-group/

Group Discussion for **5 minutes**

Use the following notes to guide the group discussion.
- Technical Requirements:
    - Data modeling and mining
    - Statistics
    - Predictive modeling and analytics
    - Tableau or Qlik for data visualization
    - Database management
    - Data security and safety
    - SQL, R, Python
    - Excel, Powerpoint, Access, PowerBI
    - *This job, like many in higher ed, is asking for a broad range of skills. This might be different than if you were working in a tech company and you ended up working in a very specific role. Higher ed, nonprofits, and startups may want a wide variety of skills in their hires. Bigger companies or government jobs may want you to show depth of knowledge or subject matter expertise in a given skill or domain. Something to keep in mind as you plan your career and build your portfolio!*
- Domain Area thoughts:
    - In higher ed, but also the "Office of Institutional Equity" so perhaps get familiar with institutional DEIA policies and federal laws about this space.
        - Asking you to help create new training programs for this purpose too
    - *Do some additional research into your area of interest. You can use your portfolio to showcase that interest and your growing subject matter expertise, too. If you're wanting to work in Fintech, for example, your portfolio might include a section where you link to articles that you find interesting. Basically, make your portfolio a space that works for you.*
- Communications / core skills:
    - Ability to work independently
    - Ability to effectively handle sensitive information with discretion
    - Work with diverse stakeholders, including employees, students, faculty
    - Write and develop reports, identify trends to inform policy
    - Analyze efficacy of educational programs, tracking and sharing participation rates
    - *To showcase these skills in a Portfolio, you'd want to make sure that your presented projects are clearly formatted, easy to understand, have tangible recommendations for stakeholders. In this particular job, you'd definitely want to assume that you had a level of technical expertise that your coworkers might not have.*

Slide 11 - Final Tips - 3 minutes

- This exercise should help you identify the types of skills you want to develop and then showcase in your portfolio. Here are some final tips to consider as you start to pick projects to work on.
- Most industries and companies are in some way concerned with sales (either to other companies or individuals), so anything involving sales data is a good place to start if you're not sure what projects to work on.
- If you're not interested in a specific domain area, think about interesting datasets that are unique or not often analyzed. For example, a lot of people have analyzed COVID infections, but far fewer data scientists have analyzed data on books checked out from the New York Public Library. Find an interesting or unexpected research question.
    - Another example of an unusual data project involving NYC pizza: https://medium.com/@seyidabs/slice-by-slice-a-visual-exploration-on-nyc-pizza-consumption-on-tableau-9dd7981734d5
- A final tip: Don't get stuck trying to craft the perfect data science portfolio. The data science field changes constantly. Keep an eye on broad industry trends and pick up new skills or tools as you go. Focus on what you think is interesting. We'll share two more online articles that will help you get started on identifying what should be in your portfolio.
    - Read more:
    - https://tripleten.com/blog/posts/9-steps-to-create-an-effective-data-science-portfolio
    - https://www.datacamp.com/blog/how-to-build-a-great-data-science-portfolio-with-examples

Slide 12 - Where should my Portfolio Live? - 3 minutes
- All right - let's switch gears. Pretend you've now completed your first project. Let's think about how to share it.
- There are a couple of options for where you can build and maintain your portfolio for free. The most common and most popular option for data scientists is GitHub Pages
    - GitHub is a free and open source platform where you can share and collaborate with others on code. You can also use your GitHub profile as a portfolio.
    - We'll share some resources on how to develop that profile in the chat and in the copy of the ppt you'll receive.
    - GitHub Pages: https://pages.github.com/
    - Tutorial: https://www.youtube.com/watch?v=D9CLhQdLp8w&ab_channel=ShawTalebi
    - Sample: https://github.com/emmabostian/developer-portfolios?tab=readme-ov-file
- LinkedIn
    - A second option is LinkedIn. This is a good option if you're very new to data science and want to have a space to share some of your introductory projects. It is restrictive, however. We'll also share in the chat a few ways to highlight your work on LinkedIn.

- - LinkedIn: https://www.linkedin.com/
  - Adding Projects: https://resumewritinglab.com/blog/linkedin-project-section-writing/
  - Adding Feature Sections: https://www.linkedin.com/pulse/linkedin-portfolio-step-by-step-guide-showcase-your-work
- Medium
  - Depending on what type of projects you're going to be working on, you might consider using Medium. This is a good option if you think you're going to be doing more writing than technical analysis. Medium will let you reach a wider and broader audience. A negative to Medium is that it's not your own website so you have less control over your content!
  - Medium: https://medium.com/
  - Tutorial: https://medium.datadriveninvestor.com/blogging-for-your-data-science-portfolio-61b07da8005d
- Own Website!
  - A fourth option would be to create your own website. The downside is that this could be a lot of work to maintain and might not be free. However, you have complete control over what it looks like and how it's structured. You will need to weigh your options here.
  - Two examples of interesting portfolio sites
  - Example 1: https://www.yan-holtz.com/
  - Example 2: https://theyvonne.com/portfolio
- Again, we will share these slides after today's meeting. We definitely encourage you to spend some time reviewing these materials.

Slide 13 - What should my Portfolio look like? - 1 minute
- So once you have your portfolio platform picked out, you'll need to establish a regular format for the materials you share online. Here are some general guidelines to follow.
- First, think about how you want to tell the story of your project. Don't just copy and paste the code, explain the context of how you worked on the project and why. Why and when did you work on this project? What interested you about it? What did you learn from the process?
  - This is an opportunity to establish your unique voice and brand online.
  - Keep the structure tight. Most people will only look at your portfolio site for between 1 and 5 minutes. Your UX is really important. The design should be intuitive.

Slide 14 - Portfolio Formatting - 3 minutes
- So let's think about the actual format. Here's what we recommend:
  - Pick a short but informative title
  - Start with a 1-2 sentence summary of the data used

- ○ Explain in 1-2 sentences what skills are showcased in this project (e.g. data cleaning, data manipulation, etc.)
- ○ Say what technical tools or skills were used to accomplish the analysis (e.g. SQL, Pandas)
- ○ As you continue your analysis, showcase that your work is focused on questions a business would probably want answered. For example, what is the sales volume of product x? Is there seasonality in sales of product y? What sales representative is really driving sales of product z? Highlight key words in bold.
- ○ Include quick visualizations of the insights as you go. Put in a graph and then write 1-2 sentences interpreting that graph. Line and bar charts are most frequently used in business. Make them colorful and easy to read.
- ○ Don't forget to use good section titles to structure your analysis into easy reading
- We will share another resource in the chat which will help you build your portfolio items' structure.
  - ○ Get more details on structure here: https://youtu.be/7_VIjcFKHZo?feature=shared&t=1131

Slide 15 - Maintaining your Portfolio - 1 minute
- And now, the trick is to keep going! Collect more projects. If you take a break from building your portfolio when you find a full-time job, set a reminder to review your portfolio every 3-6 months or so, just like you would your resume or LinkedIn.
- Ask for feedback from collaborators to make sure it reads well and is structured in an understandable way!
- Finally, expect your portfolio's focus and structure to change as you evolve your data science skills. Your professional interests may also change, so as you build your site, make sure the platform allows you to build out new tabs and sections. This is a living document just like your resume.

Slide 16 - Find a Data Science Project - 5 minutes
- If you're still feeling like you don't know how to get started, you should take a look at the Data Science Projects managed by the NEBDHub, the parent organization of the Portfolio & Network Building (PNB) Group. These free projects are built by graduate students at Columbia University.
- All of the projects are open source and can be worked on at your own pace. There is no deadline for submission. Upon receipt of your final project, the NEBDHub team will provide a Certificate of Completion for your work which can be shared on your resume or LinkedIn. Browse the NEBDHub website to find a project that matches your interests and skill level.
- Use the structure we provided in today's presentation to write up the results of your work on your Portfolio site. This should help you get started.

5 minutes for group discussion and review

Slide 17 - Final Thoughts - 2 minutes