Meeting Notes

Meeting 1: General Research and Strategy Planning

Survey Development
- Create a short survey (less than 5 questions) to ask data scientists about the tools they use in their jobs.
- Send surveys as students, maintaining a student perspective.
- Consider sending surveys regularly to gather ongoing insights.
- Utilize a QR code for easy survey access.

Web Scraping and API Access
- Explore online web-scraping and API access for job postings.
- Acknowledge job sites' limitations and restrictions on data scraping.
- Strategy for limited scraping:
  - Pull recurring amounts of limited information.
  - Scrape by region (e.g., California, DC, NC).
  - Analyze how job markets vary across regions.
  - Focus on differences in banking, healthcare, and STEM jobs by region.

Interviews and Qualitative Research
- Conduct 30-minute podcast-style interviews.
- Target at least one interview per month.
- Interview candidates:
  - Faculty members (e.g., Reggie)
  - Data scientist alumni
  - Master's students
- Engage with peers to understand what information they find valuable.
- Collaborate with student council leadership for broader insights.
- Consider brainstorming sessions with Reggie and Stephanie.

Research Priorities
- Web scraping and interviews as top priorities.
- Explore how interview processes vary between companies.
- Companies of interest: Deloitte, Capital One, etc.
- Identify top ten employers of UVA graduates.
- Understand hiring processes and expectations for data scientists.

Information Collection and Presentation
- Collect information on:
  - Company hiring processes.
  - Online channels discussing data science careers (e.g., Reddit, StackOverflow).
  - Existing surveys and their applicability to data scientists.
  - ATS (Applicant Tracking System) resume screening score calculations.
- Choose a database or cloud drive for information storage:
  - Create a Teams drive or GitHub.io page.
  - Use Jupyter Book for markdown-based website compilation.
  - Utilize GitHub Action for automated data updates.
  - Consider using PythonAnywhere for deployment.

- Communication through Microsoft Teams.
- Adopt an active-learning approach with weekly lab meetings.

Meeting 2: Career and Student Engagement (01/22)

Focus Group Discussions
- Conduct 15-minute conversations with small focus groups.
- Propose surveys during lunch events to maximize participation.
- Address the gap in students' understanding of data science applications across industries (e.g., fashion vs.
- Design surveys with open-ended questions and industry-specific prompts.

Employer Expectations
- Highlight both technical and soft skills.
- Emphasize real-world problem-solving abilities.
- Provide incentives for survey participation (e.g., lunch, raffle).

Engagement Techniques
- Utilize small group discussions to foster deeper relationships.
- Accommodate introverted participants by allowing them to write ideas on post-it notes.
- Organize post-it notes into themes and encourage group discussions.
- Implement voting mechanisms (e.g., three sticker votes per participant) to prioritize ideas.
- Apply design thinking principles to generate actionable insights.

Industry Collaboration and Internships
- Collaborate with student-led industry focus groups for interviews.
- Rethink internships for underclassmen at UVA.
- Shift the narrative around internships as a learning experience rather than an end goal for first- and second

---

Meeting 3: Advanced Research Projects and Tools (02/05)

Project Ideas and Tools
- Capstone project proposal:
  - Connect an API to an existing LLM (Large Language Model).
  - Create a RAG (Retrieval-Augmented Generation) database.
  - Implement a GPT that integrates with the FRED (Federal Reserve Economic Data) database.
  - Develop a Postgres DB with educational ML datasets for SQL queries by students.

Research Exploration
- Trial LeetCode and other technical practice websites for two weeks.
- Evaluate features and present research findings.
- Identify valuable data sources for analytics or potential research papers.

Data Collection and Permissions
- Compile job boards, analyze access costs, and assess available data.
- Contact Professor Wright to secure scraping permissions.
- Document interviews and develop blog content to share findings.

Meeting 4: Publishing and Research Tools (02/19)

Content Publication
- Set up Medium or "Towards Data Science" accounts to publish research articles.
- Experiment with a career-oriented RAG-based chatbot.

Pipeline Development
- Meet with Ali to discuss the RAG distribution pipeline.
- Establish a central repository for resources.

Research Tools and Citation Management
- Introduce Zotero for research article management:
  - Set up accounts and download software.
  - Use Zotero's extension for web link capture and citation generation.
  - Enhance organization with a shared space for research sources.

Action Items and Next Steps

Immediate Priorities
- Focus on web-scraping strategies and conducting interviews.
- Prepare white pages of information for target companies.
- Develop surveys and outreach strategies for students and employers.

Short-Term Goals (Next 2 Weeks)
- Report scraping results.
- Trial and evaluate technical practice websites.
- Draft interview summaries and publish initial blog posts.

Long-Term Goals
- Establish an experimental research lab environment.
- Maintain weekly active-learning sessions.
- Explore innovative approaches to internship experiences for underclassmen.
- Expand on data analytics projects with API integrations and GPT models.