COURSE OUTLINES: BF3228 Equity Investing with Big Data

Academic Year	2022/2023	**Semester**	2
Course Coordinator	Byoung-Hyoun Hwang		
Course Code	BF3228		
Course Title	Equity Investing with Big Data		
Pre-requisites	AB1201 and BF2219		
No of AUs	3		
Contact Hours	3hr x 13 weeks = 39		
Proposal Date	September 19, 2022		

A) Course Aims

Investment firms increasingly draw from big data as they manage their financial assets. The growing relevance of big data in the investment industry requires that students interested in a career in finance be deeply familiar with big data and its use in investments. This course aims to provide interested students with such knowledge and insight. Students will learn about the different types of big data and apply them to different investment styles. Students will learn how to program in SAS. Students will utilize their newly acquired programming skills in a series of hands-on projects with actual, real-world data and learn how to incorporate both structured and unstructured big data into their investment decisions. While the course will focus on equity investing, the knowledge and skills acquired are easily transferable to other financial assets, such as fixed-income securities and real estate.

B) Intended Learning Outcomes (ILO)/Objectives

By the end of this course, you should be able to:

1. Describe the main types of trades you can place in financial markets
2. Differentiate the main types of investment styles and analyze stocks within each style
3. Describe the main types of big data; recognize the potential as well as challenges associated with big data
4. Assess what types of big data are most useful for which investment styles
5. Analyze large datasets with SAS
6. Assemble structured data from unstructured data with SAS
7. Write a SAS computer program to evaluate a dataset's historical usefulness in predicting stock returns and develop a program that incorporates the dataset into your investment decision-making process.
8. Construct an optimal portfolio and evaluate your portfolio's performance
9. Explain the hedge fund industry and various trading strategies employed by hedge funds

C) Course Content

1. Trading in Financial Markets
2. Equity Investing – The Discretionary Approach
3. Equity Investing – The Quantitative Approach
4. Main Types of Big Data
5. Using Big Data within the Discretionary- and the Quantitative Approach
6. Optimal Portfolio Construction and Performance Evaluation
7. Hedge Funds
8. Programming in SAS
9. Creating Structured Data from Unstructured Data – Textual Analysis

D) Assessment (includes both continuous and summative assessment)

Component	ILO Tested	NBS Learning Goal	Weighting	Team / Individual	Assessment Rubrics
1. Problem Sets	ILO2-4, ILO6-9	Problem Solving and Decision Making, Motivation and Development of Self & Others	20%	Team	NA
2. Google Trend Project	ILO1-4	Creative Thinking, Critical Thinking, Planning and Execution, Oral Communication, Teamwork and Interpersonal Skills	35%, of which 25% reflect a group score and 10% reflect an individual presentation score (every member is required to present) All group members receive the same group score unless the peer evaluation indicates that the member did not contribute.	Team + Individual	Appendix 1A, 1B
3. SAS Project	ILO5-7	Creative Thinking, Critical Thinking, Planning and Execution, Written Communication, Teamwork and Interpersonal Skills	35%	Team	Appendix 1B, 2
4. Participation	ILO1-9	Critical Thinking, Oral Communication	10%	Individual	Appendix 3
Total			100%		

Description of Assessment Components:

Component 1. At the beginning of four seminars, I will distribute a problem set regarding the material covered in the previous seminar. You will have 60 minutes for each problem set. You will solve the problem set within a group. Each group comprises around six students. The problem sets will be a combination of multiple-choice questions and workout problems. The problem sets will be open book and open notes. The problem sets are not a test in a strict sense. The goal of these problem sets is to encourage you to learn from each other and provide quick feedback on whether you understand the concepts introduced in the previous seminar. To this end, I will offer loose guidance/go over the corresponding material again if I see that many of you are stuck on a given question. I will grade the

problem sets on a pass-fail basis. If your score in a given problem set is above 70%, all group members will receive a pass. For each pass, you will receive 5 points. If you pass all four problem sets, you will receive a total of 20 points. I will return the graded problem sets to you within a week.

Component 2. The Google Trend Project is a group assignment. Each group comprises around six students. I will ask you to use Google Search data to help predict a company's earnings. The final output will be your Excel sheet and an in-class presentation; each group member is required to present. Please review Appendix 1 for more details and a discussion of the assessment criteria.

Component 3. The SAS Project is a group assignment. Each group comprises around two students. In the SAS Project, you will backtest an investment strategy. The final output will be a written report. Please review Appendix 2 for more details and a discussion of the assessment criteria.

Component 4. During seminars, we will frequently solve problems together. It is thus crucial that you participate. Please review Appendix 3 for a discussion of the participation assessment criteria.

E) Formative feedback

Feedback is central to this course. In addition to your scores to the problem sets and projects, you will receive written feedback on your projects.

F) Learning and Teaching approach

Approach	How does this approach support you in achieving the learning outcomes?
Seminars	The seminars will be interactive and there will be ample opportunities for questions and open discussions. I will frequently "flip the classroom" and use our actual class time to alleviate any remaining uncertainties that you may have regarding the material; we will also appraise and deepen our understanding of the material through live problem-solving, ensuring that the targeted learning outcomes will be achieved.
Problem Sets	While the problem sets are obviously meant to evaluate your grasp of the relevant material, I will also design the problem sets to make them educational and grounded in reality so that you may further deepen your understanding of the material and learn how to apply the concepts introduced in this class to the real world.
Projects, In-Class activities	Some learning outcomes for this course are skills which are practical in nature and cannot be achieved by reading and writing alone. The achievement of such learning outcomes requires hands-on experience. The Google Trend Project, the SAS Project, as well as various in-class activities will provide such opportunities.

G) Reading and References

- Required:
  - Seminar notes, industry reports, analyst reports, news articles (posted on NTU Learn)

- Optional:
  - "A Random Walk Down Wall Street: The Time-Tested Strategy for Successful Investing Hardcover" by Burton G. Malkiel (W. W. Norton & Company; Twelfth edition (January 1, 2019))

- "Efficiently Inefficient: How Smart Money Invests and Market Prices Are Determined" by Lasse Heje Pedersen (Princeton University Press; Reprint edition (September 17, 2019))

- "The Book of Alternative Data: A Guide for Investors, Traders and Risk Managers" by Alexander Denev and Saeed Amen (Wiley; 1st edition (July 21, 2020))

H) Course Policies and Student Responsibilities

(1) General

- My goal is to frequently "flip the classroom," that is, have you review the material prior to class and then use our actual class time to clarify confusions/uncertainties you may still have regarding the material and work on live problem-solving. To this end, it is absolutely crucial that you complete all assigned readings and activities prior to class.

- I expect everyone in class to pay attention and not to distract others.

- Unless stated otherwise, kindly refrain from using laptops and other smart devices during the seminars.

- Punctuality matters so that we can start and end on time.

- In the event of a major campus emergency, course requirements, deadlines and grading percentages are subject to changes that may be necessitated by a revised semester calendar or other circumstances beyond the instructor's control.

(2) Absenteeism

- If you cannot attend a seminar for a university approved reason and there is a problem set, **inform me via email prior to the corresponding seminar**. University approved reasons include falling sick supported by a medical certificate and participation in NTU's approved activities supported by an excuse letter from the relevant bodies. I will arrange an alternate problem set taking date/time for you. You are responsible to follow up with course notes and course related announcements for seminars you have missed.

I) Academic Integrity

Good academic work depends on honesty and ethical behaviour. The quality of your work as a student relies on adhering to the principles of academic integrity and to the NTU Honour Code, a set of values shared by the whole university community. Truth, Trust and Justice are at the core of NTU's shared values.

As a student, it is important that you recognize your responsibilities in understanding and applying the principles of academic integrity in all the work you do at NTU. Not knowing what is involved in maintaining academic integrity does not excuse academic dishonesty. You need to actively equip yourself with strategies to avoid all forms of academic dishonesty, including plagiarism, academic fraud, collusion and cheating. If you are uncertain of the definitions of any of these terms, you should go to the academic integrity website for more information. Consult your instructor(s) if you need any clarification about the requirements of academic integrity in the course.

J) Course Instructors

Instructor	Office Location	Phone	Email	Consultation Hours
Byoung-Hyoun Hwang	S3 B1a - 11	67904635	bh.hwang@ntu.edu.sg	Thursdays, 1:00 pm – 2:00 pm + By prior appointment via email

K) Planned Weekly Schedule

There are three sections of BF 3228 this semester:
(1) Mondays, 9:30 am – 12:20 pm in S3-SR7/CR2,
(2) Mondays, 2:30 pm – 5:20 pm in S3-SR7/CR2,
(3) Thursdays, 9:30 am – 12:20 pm in S3-SR6/CR1.

Please attend the section you are registered to.

Week / Location	Topic	ILO	Readings*/ Activities#
Week 1 (January 9, 12) / S3-SR7, 6	Trading in Financial Markets	1	*Seminar Notes in folder "**Week 1**" on NTU Learn #Going over concepts, solving problems
Week 2 (January 16, 19) / S3-SR7, 6	The Discretionary Approach	2	*Seminar Notes in folder "**Week 2**" on NTU Learn #Going over concepts, solving problems #Valuing Tesla **#Problem Set 1**
Week 3 (January 23, 26) / **E-Learning Week (CNY)** **No consultation hours on Thursday**	Overview of Main Types of Big Data and their Possible Uses within the Discretionary Approach	2, 3, 4	*Seminar Notes in folder "**Week 3**" on NTU Learn #Going over concepts **Google Trend Project Given**
Week 4 (January 30, February 2) / S3-SR7, 6	The Quantitative Approach (Part 1)	2	*Seminar Notes in folder "**Week 4**" on NTU Learn #Going over concepts, solving problems **#Problem Set 2**
Week 5 (February 6, 9) / S3-SR7, 6	The Quantitative Approach (Part 2) and Big Data's Possible Uses within the Quantitative Approach	2, 3, 4, 8, 9	*Seminar Notes in folder "**Week 5**" on NTU Learn #Going over concepts, solving problems **#Problem Set 3**
Week 6 (February 13, 16) / S3-SR7, 6	Google Trend Project Briefs and Consultations	1-4	#Discussing Google Trend Project Status individually by group; check folder "**Week 6**" on NTU Learn for schedule
Week 7 (February 20, 23) / **TBA**	Google Trend Project Presentations	1-4	**#Google Trend Project Due**
Recess Week			

Week 8 (March 6, 9) / S3-SR7, 6	Introduction to SAS (Part 1)	5, 6	*SAS Tutorial on NTU Learn #Writing SAS programs using datasets in folder "**Week 8**" on NTU Learn
Week 9 (March 13, 16) / CR1, 2	Introduction to SAS (Part 2)	5, 6, 7	#Writing SAS programs using datasets in folder "**Week 9**" on NTU Learn **#Problem Set 4** **SAS Project Given**
Week 10 (March 20, 23) / S3-SR7, 6	Hands-on Experience with Sentiment- and Web-Traffic Data	2, 3, 4, 5, 6, 7	#Writing SAS programs using datasets in folder "**Week 10**" on NTU Learn
Week 11 (March 27, 30) / S3-SR7, 6	Hands-on Experience with Financial Market Data	2, 3, 4, 5, 7	#Writing SAS programs using datasets in folder "**Week 11**" on NTU Learn
Week 12 (April 3, 6) / CR1, 2	SAS Project Briefs and Consultations	1-9	#Discussing SAS Project Status individually with students; check folder "**Week 12**" on NTU Learn for schedule
Week 13 (April 10, 13) / CR1, 2	SAS Project Briefs and Consultations	1-9	#Discussing SAS Project Status individually with students; check folder "**Week 13**" on NTU Learn for schedule **# SAS Project Due April 17**

ANNEX A: ASSESSMENT CRITERIA

Appendix 1A: Assessment Criteria for Google Trend Project

Things you need to cover in your slides

1. List the two search terms you came up with for each of the three stocks you are trying to predict earnings for.

2. Detail your rationale for selecting those search terms. Why should they predict fundamentals in theory (and if so, which ones: revenue, earnings, or both)? Are some search terms more relevant now than five years ago?

3. Show the correlation between the changes in your search terms and the subsequent changes in revenue and/or earnings.

4. Overall, how would you assess your ability to predict your companies' fundamentals?

Group Component:

Traits	1 Below Expectations	2 Met Expectations	3 Above Expectations
Originality/ Value-Added	Students simply predict fundamentals based on company/product names.	Students show some creativity in coming up with Google Trends search terms.	Students show substantial creativity in coming up with Google Trends search terms.
Analysis	Excel sheet contains numerous mistakes, and students inadequately comment on the calculations. Students incorrectly or inadequately describe their methodology in the slides.	Excel sheet contains only minor mistakes, and students adequately comment on the calculations. Students correctly and adequately describe their methodology in the slides.	Excel sheet contains no mistakes, and students thoroughly comment on the calculations. Students' description of their methodology in the slides is very clear and comprehensive.
Use of visual aids	Visual aid is not engaging	Visual aid is standard	Visual aid is well-done and helps with presentation
Organization of the group presentation	Information is not presented in a logical sequence	Information is presented in a logical sequence with good use of visual aids to illustrate certain key points	Information is presented in a logical sequence; well designed and attractive visual aids that simplify/summarise key pointers
Overall group dynamics and delivery	Time management is poor; delivery is not smooth; students do not present themselves as a team.	Adequate time management and delivery; present themselves as a team	Good time management and smooth delivery; excellent group spirit

Individual Presentation Component:

Traits	1 Below Expectations	2 Met Expectations	3 Above Expectations
Enthusiasm/ Audience Awareness	Shows no interest in the topic presented	Occasionally shows positive feelings about the presentation	Demonstrates strong enthusiasm during the entire presentation
Delivery	Uncertain and hesitant; the student does not maintain eye contact with the audience. Mumbles, speaks too softly, or speaks in a monotone	Generally professional; occasionally establishes eye contact with the audience Adequate variation in volume and inflection	Stands poised and confident, establishes eye contact with the audience. Good variation in volume and inflection to maintain audience interest

Computation of Total Numerical Score:

To compute your total numerical group component score and your total numerical individual component score, I will take your scores for the various traits (ranging from 1 to 3) multiplied by the weights below and add them up. Your total numerical scores will thus range from 1 to 3 (with two decimal places).

Traits	Weight
Group Component	
Originality/Value-Added	40%
Analysis	25%
Use of visual aids	10%
Organization of the group presentation	15%
Overall group dynamics and delivery	10%
Individual Component	
Enthusiasm/ Audience Awareness	50%
Delivery	50%

Appendix 1B: Team Evaluation

All group members will receive the same Group Component Score **unless** the peer evaluation indicates that the member has not contributed enough to the project. The peer evaluation is conducted as follows:

Each member will rate the other team members with regards to five traits on a scale ranging from "1" through "7," with "7" denoting the most positive evaluation. For each member, I will compute the average rating across the five traits submitted by the other team members:

1. If a member's average rating is ≥ 4, the member will receive **100%** of the overall score awarded to the team assignment.
2. If a member's average rating is < 4 but ≥ 3, the member will receive **80%** of the overall score awarded to the team assignment.
3. If a member's average rating is < 3 but ≥ 2, the member will receive **50%** of the overall score awarded to the team assignment.
4. If a member's average rating is < 2, the member will receive **30%** of the overall score awarded to the team assignment.

Ratings will be kept confidential. I will only inform you if the peer evaluation indicates that you have not contributed enough to the project and how many points you lost as a result.

The following is a description of the five traits along with a description of the scales:

Traits	Performance	
Roles and Responsibility Behaves professionally and fulfills responsibilities.	**Scant** Unclear about his/her role; refuses to take a role in the group; insists on working individually, and has limited coordination or communication with others.	**Substantially Developed** Always fulfills responsibilities; performs his/her role within the group with enthusiasm; demonstrates a willingness to work collaboratively.
	Evaluation: **Scant 1 2 3 4 5 6 7 Substantially Developed**	
Communication Identifies appropriate channels to coordinate and correspond with team members.	**Scant** Modes of communication are inappropriate, causing confusion and miscommunication among team members.	**Substantially Developed** Modes of communication are appropriate; maintains timely communication and correspondence with team members.
	Evaluation: **Scant 1 2 3 4 5 6 7 Substantially Developed**	
Conflict Resolution Resolves conflicts using a variety of approaches.	**Scant** Does not recognize conflicts or is unwilling to resolve conflicts.	**Substantially Developed** Consistently resolves conflicts through facilitating open discussion and compromise.
	Evaluation: **Scant 1 2 3 4 5 6 7 Substantially Developed**	
Contributions Contributes positively; effectively utilizes one's knowledge and expertise.	**Scant** Largely disinterested in working in a group and refuses to participate;	**Substantially Developed** Actively attends and participates in all activities and articulates ideas and opinions.

<table>
<tr><td rowspan="2"></td><td>observes passively or is unwilling to share information with other team members.</td><td></td></tr>
<tr><td colspan="2">Evaluation:
Scant 1 2 3 4 5 6 7 Substantially Developed</td></tr>
<tr><td rowspan="2"><u>Relationship</u>
Maintains cooperative interaction with other team members regardless of individual /cultural differences and respects diverse perspectives.</td><td>Scant
Rarely listens to others and does not acknowledge the opinions that differ from his/her own.</td><td>Substantially Developed
Engages in respectful relationships with all other members of the team. Embraces and accepts diverse points of view without prejudice.</td></tr>
<tr><td colspan="2">Evaluation:
Scant 1 2 3 4 5 6 7 Substantially Developed</td></tr>
</table>

Appendix 2: Assessment Criteria for SAS Project

Things you need to cover in your written report

1. Detail how you constructed your long-short portfolio.

2. Detail your rationale for constructing your long-short portfolio. Put differently, why should we expect your strategy to generate abnormal returns? Why should your strategy work in theory?

3. Plot the cumulative performance of your long-short strategy (in terms of raw returns). [Hint: Export your returns from SAS into Excel and go from there]

4. What was the average annual return on your long leg? What was the average annual return on your short leg? What was the average annual return on your long-short portfolio? What were the corresponding Sharpe Ratios? Finally, what were the corresponding alphas with respect to the Market Model and the 4-Factor Model?

5. Overall, how would you assess your strategy's performance?

You will be evaluated largely based on the following criteria:

Traits	1 Below Expectations	2 Met Expectations	3 Above Expectations
Apply Concepts to Real-World Situations	Student cannot apply the concepts learned in class to real-world situations.	Student can mostly apply the concepts learned in class to real-world situations.	Student is fully and expertly able to apply the concepts learned in class to real-world situations and expand upon them.
Analysis	SAS code contains numerous mistakes and is not adequately commented on. Students incorrectly compute/describe their portfolio construction/portfolio performance.	SAS code contains only minor mistakes and is adequately commented. Students correctly compute/describe their portfolio construction/portfolio performance.	SAS code contains no mistakes and is thoroughly commented on. Students correctly compute/describe their portfolio construction/portfolio performance.
Writing	The report is difficult to understand. Problems with sentence structure, leaving the reader unsure of the meaning.	The report can be adequately understood.	The report is well-organized, easily understood, concise, and to the point.

You will not be evaluated based on the performance of your trading strategy. If the idea behind your strategy is reasonable, you coded everything correctly and you properly addressed the five questions in the Written Report Template, you are all set.

Team Evaluation

All group members will receive the same Group Component Score **unless** the peer evaluation indicates that the member has not contributed enough to the project. The peer evaluation is conducted in the same manner as described in Appendix 1B.

Computation of Total Numerical Score:

To compute your total numerical score for the SAS Project, I will take your scores for the various traits (ranging from 1 to 3) multiplied by the weights below and add them up. Your total numerical score will thus range from 1 to 3 (with two decimal places).

Traits	Weight
Apply Concepts to Real-World Situations	25%
Analysis	40%
Writing	35%

Appendix 3: Assessment Criteria for Participation

Traits	1 Below Expectations	2 Met Expectations	3 Above Expectations
Contribution frequency	Does not speak up/ contribute in class	Occasionally speaks up/ contributes in class	Regularly speaks up/ contributes in classes
Contribution quality	No contributions/ contributions lack substance	Contributions demonstrate basic understanding of subject matter	Contributions are constructive and insightful

Computation of Total Numerical Score:

To compute your total numerical score for participation, I will take your scores for the various traits (ranging from 1 to 3) multiplied by the weights below and add them up. Your total numerical score will thus range from 1 to 3 (with two decimal places).

Traits	Weight
Contribution frequency	40%
Contribution quality	60%