Maastricht University

Syllabus

Economics and Management of Information

Network Economics

EBC 2109

Academic Year: **2021-2022**

Course Period: **5**

SBE Emergency Number: 1333

!Disclaimer! The information provided on code.unimaas.nl about the course was set up prior to the coronavirus crisis. As a consequence of the crisis, course coordinators may be forced to change the teaching and assessment methods used. The most up-to-date information about the teaching/assessment method(s) will be available in this course syllabus.

School of Business and Economics

Bachelor *Master*

Table of Contents

Introduction

The new economy is driven by dramatic changes in information and communication technologies — the computer, telecommunication technologies, the internet … These technologies, which are being integrated into every facet of the economy and society, are inherently network technologies. They gain their value from being interconnected. This course studies the micro-economics of networks, and the effects on the economy of the widespread diffusion of network technologies. But it is broader than that. The programme Economics and Management of Information focuses in part on the role of knowledge and information in the economy. Part of the role of knowledge and information is to create and diffuse innovations. But knowledge has strange properties as an economic good. It is not like apples and oranges in many ways. This course briefly discusses the issue of codified and tacit knowledge, and the importance of tacit knowledge (knowledge that can't be "written down") in innovation and growth. Knowledge that cannot be written down, or "codified" is difficult to transmit — imagine trying to write a "How to ride a bicycle in 7 easy steps" book. You couldn't do it in a way that would make it possible for a reader to ride a bike at the end of it. So if tacit knowledge matters (and we will argue that it does) how does it diffuse? Then answer is that it diffuses over social networks. This is the basic insight that guides this course. To understand knowledge diffusion, and so knowledge creation or innovation, you have to understand how networks of knowledge actors work. Thinking in terms of networks implies thinking not only about agents, resources and objective functions, but also about which agents are connected and how they are connected to each other.

The idea that agents are connected to specific other agents, for a variety of reasons, might change the way we think about how economies work. For this we need to do some network analysis (as opposed to simply observing network features of some technologies). The course introduces social network analysis, and develops it in the context of innovation, knowledge creation and knowledge diffusion. We talk about different network structures, how to measure and describe them, and the implications of different structures for the innovation performance of an economy.

Prerequisites An intermediate level of economics is recommended. Basic statistics and econometrics are needed. Exchange students should have a basic knowledge of microeconomics, statistics and econometrics, and an advanced level of English

This course is open to students enrolled in the Bachelor Economics and Business Economics - Economics and Management of Information Programme. It is also open to SBE Exchange Bachelor students.

Instructors

Course Coordinator

Prof. Robin Cowan

Boschstraat 24
Maastricht
r.cowan@maastrichtuniversity.nl

Tutor

Tony Hung

Tongersestraat 53 K. A0.02
Maastricht
tony.hung@maastrichtuniversity.nl

Course Objectives

1. In this course students will learn the basics of social network analysis. This includes understanding how to describe and measure certain aspects of networks generally, and how to interpret network statistics.
2. Students will learn about basic network structures that are commonly observed.
3. Students will learn the basic concepts of social capital, and how these are related to social networks, and to innovation.
4. Students learn to apply these concepts, both social capital concepts and network concepts in the context of innovation: knowledge production and knowledge diffusion.
5. Students will also learn the basics of the software package R, and its network analysis package *igraph*.

Course structure

The course will be organized as a modified PBL course. There will be one "standard" tutorial per week and one session in which students will work in small groups. The goal of the tutorials is to introduce basic concepts in network analysis. The goal of the small group sessions is to work through a research project together. This means that attendance is very important, as the learning is done through in-class discussion and pre-class preparation. In principle all meetings are mandatory.

There will be an introductory, plenary lecture. The lecture will be recorded so students can view it at their own convenience, however, it should be watched before the first tutorial meeting.

There are roughly two meetings per week, (depending on holidays). In principle one meeting will be a relatively standard PBL style tutorial, the other will be a meeting wherein students work in small groups. The tutorial topics and tasks for the small group sessions are laid out below. Both tutorial sessions and small group sessions demand before-class preparation. So it is important to read the syllabus for each session several days before the session takes place so you can be well-prepared when you arrive. If you are not prepared, you will not have a successful session. This is not only true for tutorials (as you should be aware, having had PBL before), but also the case for the small group sessions.

The first lecture will introduce the course and give a very general background and motivation.

Literature

At the end of this document is a literature list.

Schedule

1	Mon. 11 Apr.	Lecture R. Cowan	Introduction Available on Canvas	
1	Thurs. 14 Apr.	Tutorial	T1 What is a network T2 Networks in the economy	
2	Thurs. 21 Apr.	Small Group	SG1 Find a topic of interest and start looking for data Check R installation.	
3	Mon. 25 Apr.	Tutorial	T3 Software introduction R and igraph	
3	Thurs. 28 Apr.	Tutorial	T4 Describing networks	
4	Mon. 2 May	Small Group	SG2 Finish getting data, start thinking about research question; do you need more data?	Research question and data sources and variables. **Due Friday May 6, 23:59**
5	Mon. 9 May	Tutorial	T5 Why do networks form T6 Weak links and social capital	
5	Thurs. 12 May	Small Group	SG4 Describing the network Basic statistics, basic diagram	Table of descriptives, preliminary plot **Due Tuesday May 17, 23:59**
6	Mon. 16 May	Small Group	SG5 Make analysis plan; start on it if there is time.	Analysis plan **Due Tuesday May 24, 23:59**
6	Thurs. 19 May	Tutorial	T7 Small worlds	
7	Mon 23 May	Tutorial	T8 Scale free networks	
8	Mon. 30 May	Tutorial	T9 Example papers	
8	Thurs. 2 June.	Small Group	Final session for questions	
	Friday 10 June	Individual	Final paper	Final Paper **Due Friday June 10, 23:59**

Software

We will introduce some software in this course. We will look at the R package called igraph. R is a very general purpose statistical package and scripting language. It has several packages or libraries devoted to network analysis. The advantage of R is that it is very flexible and does not only network calculations but also any other kind of analysis you might want to do.[1]

You should install R (it is open source). And in R install the igraph package.

You should do both of these things a few days before the first small group meeting on April 21 so you can make sure they are installed and work properly.

You can download R here:

https://www.r-project.org/

A nice interface for R that many people like is called RStudio:

https://www.rstudio.com/products/rstudio/download/

R tutorials abound on the internet. Here is one that seems like a pretty good introduction. It presents R in small bites.

http://www.r-tutor.com/r-introduction

Before class you should work through the first several pages of that introduction tutorial.

About installing packages, you can look here:

https://www.dummies.com/programming/r/how-to-install-load-and-unload-packages-in-r/

Igraph tutorials also exist: https://kateto.net/networks-r-igraph is very thorough, and covers a lot of ground in a pretty accessible way. It starts with a recap of R basics and then turns to igraph. There is quite a lot about creating a graph/network in igraph, which you can probably skip at this point. In fact, there may be quite a few things here that you don't care about too much, but there is plenty that you do (for example there is quite a lot about customizing the way a network is presented as a picture).

You should look at this tutorial before class.

[1] Gephi is another network visualization tool, which will produce basic network statistics that you can export to other software to analyse. The advantage of Gephi is that it is all menu driven and has a relatively friendly interface. The disadvantage is that it is not so easy to link it with other software. Gephi: https://gephi.org/

There is a third common network software package called UCINET which you are welcome to use, but we won't be talking about it in class. Partly because it is expensive.

Assessment and Grading

There are four parts to the final grade.
From three of the small group sessions there should be a tangible output. Usually these are short pieces of texts, or a table. They are to be handed in as a group assignment. The fourth part of the grade comes from a final paper which is to be individual work. The paper is built on these group assignments, but written individually. The final paper will be due on the last day of the exam week.

The weight for each assignment is as follows:

Research question and data sources and variables (SG2)	15%
Table of descriptive statistics (SG3)	20%
Analysis plan (SG4)	20%
Final paper	45%

To pass the course you must achieve a final grade of 5.5/10 in aggregate. There is no passing requirement for individual parts of the assessment.

Due dates of all assignments will be posted in Canvas.

Overview

	Course Objectives				
Assessment	*1*	*2*	*3*	*4*	*5*
SG2: Research Question Assessment method: 1 page deliverable				X	
SG3: Describing the Network Assessment method: 1-2 page(s) deliverable	X	X			X
SG4: Analysis Plan Assessment method: 2 pages deliverable	X	X	X	X	
Final Assignment Assessment method: min. 3500 word document	X	X	X	X	X

Details of the group assignments are given in the detailed description of the tasks and schedule.

The final, individual assignment is a written paper, of roughly 10 pages including tables and figures (roughly 2500-3000 words). It will be hard to do a good job with fewer than 2000 words. It should be a research paper that builds on, and incorporates, the work you have done with your small group. In this paper you should introduce your topic, explaining why it is interesting and important. You should include a short literature review showing the literature you have read to understand the topic: roughly what it says; what conclusions it draws, if any; any gaps you see in the literature you have read. (Obviously for a course like this you cannot read every relevant paper ever written. But you should try to get a reasonable coverage using 4 or 5 (or more if you have time) papers.) After you describe your data, giving descriptive statistics, and perhaps a plot of your network (again using what you have done in your group), you present the analysis. Then finally your conclusions. In the conclusions you should discuss what you found (which might or might not support your hypotheses), implications or importance of it, and the limitations of your work or your data. One way to approach the last thing is to say, "If I had more time and more resources I would have done …".

It is important to note that in this assignment we are less interested in the results per se (in this amount of time it is very difficult to get results that are robust to all the sorts of complaints that economists love to bring up). So, we are not going to be looking for big R^2 or t-statistics. What we are interested in is the logic of what you did to try to find an explanation for or understanding of the phenomena you are interested in. At the end of this document, you will find a rubric for the final assignment, which explains in detail how to format the document and what the objectives and expectations are for it.

Resit

No resists are possible for the group assignments.

The final paper can be re-written if the student achieves a failing grade. This however does not apply if the failing grade is caused by plagiarism. If there is plagiarism, there will be no resit.

Late assignments will not be accepted without proper reasons of which we are informed before the fact.

Validity of partial results

The rule from the EER is the following:

Parts of an educational unit which were passed will remain valid in the academic year in which the partial results are obtained and two (2) more academic years, without prejudice to the competency of the Board of Examiners to extend this period of validity. In case the course changes substantially from the previous year, transitional rules for repeat students need to be clearly indicated in the course manual. Also ensure to check the distinction between 'exam only' registrations and regular registrations.

Task 1. What is a network?

"Network" is a relatively modern word in secular English, at least as a common word. It goes back to 1560, but you don't find it in the Concise Oxford English Dictionary as recently as the fourth edition (1951). Nonetheless, looking back it is possible to identify historical examples of what we would now refer to as networks. Before we move too far into the analysis of networks in the context of economics, it is worth exploring just what we mean by "network". What is a network? What is a network technology? What characteristics do networks have, and what characteristics do network technologies have? To begin this discussion, find a definition of "network"— this can be from a dictionary, from a web page on networks or (almost) anywhere else. (The mathematical field called graph theory also talks about networks.) Find two examples of networks or network technologies — one current example, and one historical (existing before 1950). These can be economic networks, technology networks, social networks or any other kind of network you can think of (but don't use the internet as one of your examples). Be prepared to discuss these two examples in the context of your definition of network. How are your two examples similar; how are they different from each other? Given the examples presented, is there a general characterization of a network?

Questions:

What is a Network?

What is a network technology?

Find two examples of networks or network technologies — one current example, and one historical (existing before 1950).

How are your two examples similar? How are they different from each other?

Is there a general characterization of a network?

Task 2. Networks in the economy

You should now have a pretty good idea about networks. The claim will be made that analysing economic phenomena using network analysis, or using network-based models is a good thing to do. It will advance our understanding of how markets work. But first, can networks be used to analyse any market or economic phenomenon? The answer is clearly "Yes. You can use any tool you like to analyse any phenomenon you like. It is just that some analyses will be good ones and some will be bad ones." In other words, network analysis might be good for looking at some phenomena, and bad for looking at others.

In economics, when seen as a discipline as a whole, the models used to analyse things do not traditionally resemble networks very much. Consider, for example, the standard partial equilibrium model of perfect competition, or its big brother, the general equilibrium model.

How are they "un-network-like"? (I don't mean that you have to write one down — just think about the models you have seen in the course of your studies, and the properties they have.) Now, starting with the answer to that question, you should be able to compile a list of features that a network model will have. Of course, not every model that calls itself a network model will have every single feature, but there should be some common, central elements.

I just said that any tool can be used to analyse any phenomenon if you like, but some tool-phenomenon combinations work better than others. You should now have a reasonably good idea of how network and non-network tools differ (and are similar), now think about how phenomena could be grouped according to whether they are likely to be usefully analyzed using network-type tools or not. Can you see markets/industries/technologies/ goods that should be modelled as networks and others that cannot or should not?

Questions:

What are the properties of classic economics models (like General Equilibrium Theory)?

What are the properties of network models?

How do the two types of models differ?

Find markets/industries/technologies/goods that should be modelled as networks and others that cannot.

Before the first small group session you should read Task 3 and install the software that is described there. Instructions are pretty clear. You will check your installations during the first small group session.

Small Group Session 1. Topic and Data

In this session your group should discuss what could be an interesting phenomenon to investigate. You should find a topic of common interest that could lend itself to network analysis. This means that to understand something you observe, it helps to think in terms of networks. This is more than simply describing a network. It is usually about using network ideas to understand something else.

Having found a topic, you should look on the internet for data that you could use to investigate it.

In the course of this discussion you should think about what you consider a node in the network, and what it is that links two nodes that you think are connected.

Some examples:

Does the centrality of a student in the friendship network predict his or hers grades?

Does the network of regional trade agreements predict trade flows?

Does the transportation network predict the spread of covid or SARS or Ebola or…?

Does centrality in twitter affect re-tweets?

Facebook provides connection data at the county level. Is the position of a county in this network correlated with its income or wealth (or other measure of welfare)?

Is position in a friendship network correlated with the probability of finding a (good) job?

Modern rapper network and performance.

Tourism and spread of covid-19

Co-authorship and citations

Development of Indian railway network and prosperity

Network structure and adolescent alcohol consumption

Road network and local GDP

Board interlocks and innovation/sales/… performance of some kind

S Senate bill co-sponsorships and the success of bills

Movie franchise networks (like Harry Potter) --- can they tell you anything about the plot or audience popularity ?

There are several standard sources for network data:

https://github.com/briatte/awesome-network-analysis (this is an aggregation site)

http://vlado.fmf.uni-lj.si/pub/networks/data/UciNet/UciData.htm;
http://www-personal.umich.edu/~mejn/netdata/;
http://snap.stanford.edu/data/

- jazz musicians http://deim.urv.cat/~aarenas/data/welcome.htm
- network Science co-authorship
  http://www-personal.umich.edu/~mejn/netdata/
- Florentine families of the 15th century
  http://www.casos.cs.cmu.edu/computational_tools/datasets/external/padgett/index2.html

(available in many places, also available as an R dataset)
Interlocking Corporate Boards (Stokman-Ziegler)
http://vlado.fmf.uni-lj.si/pub/networks/data/UciNet/UciData.htm#konec

If you use the Chrome browser (or a derivative of it) you can install the "Lost Circles Social Network" extension. It will download and draw the network among your Facebook friends (assuming you have a Facebook account). You can then download that graph data (use GraphML format) and open it in igraph (or Gephi).

People also work on trade data using network analysis

- In this session you should also check your R installation. (See the text for Task 3.)

Task 3. R and igraph

This will be a session on two pieces of software: the statistical package called R, and a library called igraph. The former is multi-purpose software for statistical analysis and much more. The latter is a library for R which facilitates network analysis.

R is now very popular, and getting more so. In fact in job advertisements one often sees that knowledge of R is considered an asset.

Learning R and igraph can be challenging sometimes. The key is to use the internet a lot — everything you try to do someone else has already done, and someone has asked the internet how to do it. So be prepared to have your browser open to your favourite search engine while working in R.

You can download R here: https://www.r-project.org.

There is an interface that many people like to use called RStudio. It is available here:

https://rstudio.com

This looks commercial, but there is a free version which will do everything you need.

R tutorials abound on the internet. Here is one that seems like a pretty good introduction. It presents R in small bites.

http://www.r-tutor.com/r-introduction

Before class you should work through the first several pages of that introduction tutorial.

About installing packages, you can look here:

https://www.dummies.com/programming/r/how-to-install-load-and-unload-packages-in-r/

Igraph tutorials also exist: https://kateto.net/networks-r-igraph is very thorough, and covers a lot of ground in a pretty accessible way. It starts with a recap of R basics and then turns to igraph. There is quite a lot about creating a graph/network in igraph, which you can probably skip at this point. In fact, there may be quite a few things here that you don't care about too much, but there is plenty that you do (for example there is quite a lot about customizing the way a network is presented as a picture, and how to calculate network statistics).

You should look at this tutorial before class.

We will inform you of the logitstics of this meeting closer to the time.

Small Group Session 2. Research Question

In this session you should finalize your research question. What you are going to study, and what do you hope to find out. With that set, you need to look again at your data to ensure that you can address your topic. If you need more data, start to find it.

This session comes with an assignment that must be handed in. It is a short document, about one page.

1. Begin with a short paragraph stating the general phenomenon you are looking at, and a few sentences about why it is interesting and important.

2. Then your carefully worded research question, or statement what it is you want to find out with your study.

3. Then a list of the data you will use: the dataset, the place you found it, what are the variables in the data that you will use.

4. State clearly in your analysis what will be a node, and what will be a link between nodes.

Each of the four elements listed above will receive equal weight in the grade.

On the topic of research questions, here are two places to look:

https://writingcenter.gmu.edu/guides/how-to-write-a-research-question

https://owl.purdue.edu/owl/general_writing/common_writing_assignments/research_papers/generating_questions_and_topics_workshop.html

Task 4. Describing networks

A network is a population of agents and links between pairs of agents. Different networks have different structures. What could that mean? How would we describe a network? One way is just to list who is connected to whom.

a) Networks can be "weighted" or "unweighted"; "directed" or "undirected". What do those mean?
b) Suppose you have a population of 2 people. How many possible networks are there? (For the purposes of part b) think only about unweighted networks.) This you can do by drawing them. Now what about 3 people — how many possible networks are there? Now generalize. Suppose there are n people. How many possible networks are there? What if $n = 100$? Does it matter if we are talking about directed or undirected?
c) Part b) should tell you that describing networks using a list of who is connected to whom is not going to be very useful. So we need something else. Find the definitions of these network statistics: degree; density; clustering (which is the same as cliquishness or transitivity); betweenness; path length; average path length. What is a component?

Questions:

Know the basic network statistics from the reading.

Can you calculate how many possible links there are if there are n nodes?

Can you calculate how many possible networks there are if there are n nodes?

How do the networks explored in your group differ from each other? How are they similar? Why?

Task 5. Why do networks form?

A network consists of a collection of agents, some of whom are connected to each other. Further, we suggested that for the network idea to be interesting for economics these connections should have some temporal duration, that is, they should last for some significant length of time. The task here is to answer "Why?" Why do networks form, and why should these connections have significant temporal durations? This is a big question and there are many things that can contribute to the answer, that is, there are many reasons to form connections with other agents, and many reasons to form connections with more than one other agent (since a group of pairs of agents is not really that interesting as a network).

a) Start simply by asking this: Economists think that market transactions have lots of really nice properties both at the micro level and at the macro level. But we observe firms forming non-market connections with each other. One form of such a connection is the strategic alliance. Why would firms form strategic alliances rather than simply engage in market transactions?

b) One way to think about network formation is to take a particular type of network: some networks are about moving goods — inputs and outputs — around (firms connected to firms); some networks are about moving knowledge around. In the context of the latter, an idea called "tacit knowledge" has become quite important in understanding networks. Tacit knowledge is contrasted with codified knowledge. Here is a claim: "If all knowledge is completely and effectively codified, then networks are unnecessary for the transmission of knowledge." What could this mean? It is connected to the next paragraph.

c) Another thing that we should note is that in a sense networks in which every agent is connected directly to every other agent are not so interesting either. (In fact they are a lot like a perfect market, so for an economist, network theory may not add that much insight if everyone is connected to everyone else since we already have good models of that situation.) So why do I connect only to a small number of other agents? If it really is the case that typically any agent will connect with only a small number of other agents (small is relative to the total number of agents in the economy), what does this tell us about

network architecture, if anything. (By architecture I mean who is connected to whom. If you drew a map of connections, what would it look like — would it be totally random or would there be an identifiable pattern? If the latter, what would the pattern look like.) In the abstract, it tells us nothing — that is, the pattern would be random since anyone could in principle be connected to anyone else located anywhere else. But when you think of the reasons that people only connect to a few others, would you expect some pattern to emerge? (A classic reference here is to Alfred Marshall and "Marshallian externalities".)

Some key words to search on: "transportation costs"; "tacit knowledge"; "sticky data" or "sticky knowledge" "Marshallian externalities".

Questions:

Define tacit and codified knowledge.

If all knowledge is completely and effectively codified, then networks are unnecessary for the transmission of knowledge. What could this mean?

So why do I connect only to a small number of other agents? What does this tell us about network architecture?

When you think of the reasons that people only connect to a few others, would you expect some pattern to emerge?

Task 6. Strength of Weak Ties, Social Capital and related issues

In previous tasks we discussed network formation. If networks form, the agents forming them must be getting some benefit from doing so (otherwise why would they go to the trouble?). This benefit is referred to as "social capital". There are a large number of definitions of "social capital". Here are three, associated with prominent sociologists.

- "the aggregate of the actual or potential resources which are linked to possession of a durable network of more or less institutionalized relationships of mutual acquaintance and recognition." Pierre Bourdieu
- "a variety of entities with two elements in common: they all consist of some aspect of social structure, and they facilitate certain actions of actors...within the structure." James Coleman
- Social capital "refers to the collective value of all 'social networks' and the inclinations that arise from these networks to do things for each other." Robert Putnam

Read these definitions, and any others you find. What is common to them? What is different? How do the differences (if there are any) matter?

One question that immediately arises when thinking about social capital and networks is where precisely social capital comes from. Does it arise in some specific of the network structure, perhaps combined with properties of the nodes?

Mark Granovetter believes that social capital arises in "weak ties".

James Coleman believes that social capital arises from "closed positions".

Ronald Burt believes that social capital arises from "structural holes" (or "open positions").

So your first task is to figure out what those things mean.

One way to think about this is that network ties provide opportunities but also impose constraints. How do those combine, within different local network structures, to provide benefits for the agents involved? Here you should relate the ideas of Granovetter, Coleman and Burt. Do these ideas overlap? Are there similarities? How do they differ?

Then: There are three different hypotheses here (Granovetter's, Coleman's and Burt's). For each hypothesis, under what conditions would it be true, under what conditions would it be false?

Part b) When instead of just thinking of nodes as nodes, but rather as agents that have properties, it could be that agent properties affect network formation and the opportunities and constraints that networks provide. In this regard, a prominent concept is "homophily".

What does that mean?

The death of distance has been much discussed. Modern information and communication technologies remove the constraints of geography, or so it is claimed. But geographical distance is only one sort of distance. What other types of distance are there? Do they matter regarding the ability of agents (people, firms, institutions) to communicate or collaborate successfully? Notice that when you define "distance" in a very general way, you can use the concept of homophily equally broadly.

Some people (Brynjolfson and van Alstyne on the Balkanization of Science for example) claim that in fact they make it *too* easy for people to collaborate over long distances. Why would they claim that?

Two concepts can be important here: homophily and strong/weak ties. Figure out what they mean.

Homophilous ties, or strong ties, have good and bad effects. What are they? One place you can look is at the literature on "referral hiring" and "occupational segregation". This is strongly connected to homophily.

Questions:

What properties does social capital have?

Define "weak ties", "closure" or "closed network position", "structural holes".

Understand how and under what conditions social capital could be generated by each of those different network structures.

Define “homophily”.

Explain how homophily might affect network formation.

Describe some of the effects of homophily on how a network performs and what implications it has on, for example, science, hiring.

Small Group Session 3. Describing the network

Now you will use the things you have learned in tasks T3 and T4 to start describing your data and your network. This session comes with an assignment. Again, it is a short document, mostly consisting of a couple of tables.

1. Your document should begin with a very short paragraph describing your research issue, and the data you have. State what are the nodes and what are the links.

2. Then a table of statistics that describes the network. These are the (relevant) statistics you discussed in task 3 (things like average degree, average path length and so on). If there is anything striking about the network you should point it out.

3. Probably you are connecting your network to some other data. What are the variables that matter in those data? Put them in a table also.

In both tables, if you are dealing with node-level data, usually it is relevant to give mean and standard deviation of the statistics.

4. It is also standard in this section of a paper to include a correlation matrix that shows which variables are correlated with each other.[2]

Each of the four elements listed above will receive equal weight in the grade.

Bonus points if you include a “beautiful picture” of the network. Draw it using igraph, and try to draw it in a way that illustrates some of the features you think are important. You might add a short paragraph pointing them out.

This assignment should be a page or two long, depending on how many variables or statistics are relevant.

[2] If your variables are in vectors called x, y, and z, you can use the R command cor(cbind(x,y,z)) to get the correlation matrix easily. If your variables are already in the columns of a matrix (called M) or a dataframe, it is even easier: cor(M) does it.

Small Group Session 4. Analysis plan

The goal here is to discuss how you should analyse your data. You have an issue, you have some data, and you have enough descriptive statistics to know what you have got. The issue for today is what to do with all this stuff to make an argument that addresses your issue or answers your research question.

Again you need to submit a document. Again not more than 2 pages, probably less. State your question or issue briefly (remember we have seen your other assignments so this is more to remind us); describe (again briefly) your data. Then say what you will do in order to analyze your data to address your research issue. What we are interested here is whether you can form a coherent strategy that will address your issue, and try to control for factors that you are not interested in.

Below is a template you can use to prepare your analysis plan:

Analysis Plan Template	
Date:	**Name of the project:**
Team Members:	
Introduction to the topic and Research Question	
Brief introduction:	
Research Question:	
Hypothesis:	
Data details	
Data set(s) used:	**Analysis Software:**
Dependent variable(s): (what you are trying to explain)	
Independent Variables: (what you think would explain it)	
Proposed Analytical Strategy	
Data Manipulation: Do you need to manipulate your data? For example, you would like to use per capita income as variable, but you don't have it in the data set. However, you do have variables with household income and household size, so you can use those two to generate your variable of interest.	

Data Visualization: if you draw a scatterplot or a boxplot, can you spot a pattern?
Data Analysis: Will you use multivariate regressions, correlations, ANOVA, ANCOVA, etc.? What graphs or tables will you include?

Your grade will depend on:

1. Your ability to fill in the template correctly..
2. The generation of an adequate hypothesis
3. Identification of the dependent and independent variables
4. Techniques you will use to analyse your data

Task 7. Small worlds

In the previous task you were asked about whether the reasons for which and the mechanisms by which agents form links in a network would create identifiable structures in the network. In this and the next task we look at two such structures.

It is a pretty common experience to meet someone who knows someone you know. We say "What a small world it is."

Have a look at this web site:

http://oracleofbacon.org/ and play the Kevin Bacon Game. How well do you do? Why do you do well or badly? What is the Kevin Bacon Game, and what does it illustrate? (Barabasi in his book *Linked* talks about the origin of the game. It started with bored students.)

A small world has the features that in general an agent is connected to only a small number of other agents; there tends to be a lot of overlap in these connections (my friends are likely to be friends of each other) but nonetheless, the distance between two agents (calculated as the number of steps between them) is relatively short. Modern, network interest in small worlds began with an experiment in 1967 by Stanley Milgram (on the reading list). What did he find? There are modern versions of this experiment. Find one and report on it.

Duncan Watts and others claim that small worlds are very common, and that they tend to be a pretty common structure for all kinds of things. Understand what this means, and why it is true. A good place to start is Cowan and Jonard, Network Structure and the Diffusion of Knowledge.

1) Part of the key here is to understand cliquishness and path length, and what each of them does. 2) Then ask why high cliquishness is good, and why short paths are good in the context of economics.

3) Finally, ask whether there are conditions under which these could be "bad things".

Questions:

What is a Small World network? Describe the model

Why are small worlds a common structure?

Find a modern version of Granovetters experiment and report on it.

Part of the key here is to understand cliquishness and path length, and what each of them does.

Then ask why high cliquishness is good, and why short paths are good in the context of economics.

Ask whether there are conditions under which these could be bad things.

Finally, how does this discussion relate to the concepts in Task 6. ?

Task 8. Scale Free Networks

If you start by thinking about agents, you can see that networks emerge because agents form links with each other. The rules they use for forming links (pick another person at random; pick the closest available person; pick a person with such and such characteristics …) obviously determine the structure of the network that emerges. The simplest to understand is the random network.

The mathematical tool used to study networks has, until recently anyway, been something called graph theory. The big breakthrough came in the1960s when Paul Erdos (see task **9**) and Alfred Renyi laid the foundations of random graph theory. This approach held sway for a long time partly because there were nice theorems that could be proved about random graphs. The basic idea is that the way to think about networks is as if we had a large number of nodes and they are linked by a certain number of "edges" (connections). The key is that the edges are allocated randomly among the nodes. So if you think of building a network, you start with a fixed, unchanging set of nodes (agents), and then over and over again choose two nodes at random and put an edge between them. Stop when you have the desired number of edges in total. Now you have a network. Sometimes this describes how networks form. But even if it doesn't describe the actual formation mechanism of a network, it might result in a network that looks like some real network.

Questions: If you do this, what does the resulting network look like? (Think about drawing a picture of it—what would that picture look like?) Next, what does the frequency

distribution of edges look like? (ie, how many nodes have 1 edge? how many have 2 edges? etc.)

Second question: What are some examples of networks that form like this; and/or networks that don't necessarily form like this but which have this sort of final structure?

And the more interesting question perhaps: Examples of networks that do not look like this, and just as important, what do they look like?

You should be able to come up with examples just by thinking hard about the networks you are familiar with. Use your imagination, as lots of different things can be characterized as networks. For example, think about social networking. What is the distribution of "friends" in Facebook (ie how many people have 0 friends, how many have 1 friend…how many have 1000000 friends.) Similarly with twitter followers. Or Facebook likes. Do these look like the degree distributions of a random graph?

Among your facebook friends, what is the degree distribution (you should be able to get this information either by asking, or by looking at their pages). Does it look like the overall distribution? Why or why not?

If you use the Chrome browser (or a derivative of it such as Iridium, Opera or Vivaldi) you can install the "Lost Circles Social Network" extension. It will download and draw the network among your Facebook friends (assuming you have a Facebook account). (See comments in task 7.) You can then download that graph data (use GraphML format) and open it in Gephi or igraph. Then you can generate network statistics to describe that network, and ask questions like "Is it a small world?" "What is the link distribution?" "Who is the best connected?"

There is one type of network which does not look like the random network, but that seems to be pretty common, and this is called the scale-free network. This is a network whose frequency distribution is described by a power law: The probability of a node having *k* edges or links, or equivalently (for large networks) the proportion of nodes having k edges, is $P(k)=Ak^b$, where *A* is just a scaling factor, and *b* is the scaling degree or exponent, and is greater than 1. So if you ask, how many nodes have 1 edge, the answer is $P(1)=A$. How many nodes have 2 edges? $P(2)=A \times 2^b$. How many have 3 edges? $P(3)=A \times 3^b$. And so on. (Draw something like this, and figure out what it means. I suggest you use a spreadsheet and fool around with *A* and *b* to see what effects they have). Here is the tough question: What algorithm creates this kind of network? It is easiest to see in a growing network. Start with one node, and then add nodes. But each time a node is added it must be connected. Suppose that when you add a node you connect it with *d* edges (where *d* is the same for every node you add—some number like 2 or 3). If you just pick nodes at random to connect your new node to, will you get the random graph of Erdos and Renyi? (The answer is no; why not?) But you don't get a power law either. So how do you have to add

them to get a scale free network? And of course the question: are there networks that are formed like this? Why? What sorts of networks would be formed this way?

Finally, what are the differences in performance between a scale-free and a random network? Think about a few things—suppose you are trying to diffuse information across networks; suppose you are trying to stop information from being diffused for example. In place of "information", you could put "disease", or many other things. Or suppose you are trying to disrupt airline traffic.

Questions:

What is a Scale-Free Network?

What is an Erdos-Renyi Random Network?

What does the distribution of the edges look like in a Random Network?

What are examples of random networks and of networks that does not necessarily form like this but that have this kind of structure?

Find a definition for a scale free network.

What algorithm creates a scale free network?

Task 9. Example papers

In this task the point is to look at some examples of how people have used network analysis to address different issues. The literature list contains several papers. Each sub-group of students will select one paper and present it to the rest of the group. A presentation will be 10 to 15 minutes long, which should leave some time for questions, given the number of student per tutorial.

What should go in the presentation?

- A statement of the empirical context.
- A statement of the research question the paper asks, or the issue or phenomenon the paper is trying to address.
- Brief description of the data used (or created).
- What is the argument? Here you should be referring to the concepts you have been discussing throughout the class. This is the key part of the presentation. How do those concepts help us to understand the phenomenon the paper is addressing?

Final Assignment Rubric

Objective of the final paper:

Students are expected to put together a research document with a typical academic format, including all the elements specified in the provided structure below. They should be able to implement methods and concepts covered during the course and correctly interpret their findings both in terms of the direct measured output as well as implications for the larger economic context.

Students are also expected to write a coherent document that synthesizes their acquired knowledge throughout the course and supplement that information from new academic sources.

The final assignment is due June 10, 23:59.

Structure:

1. Introduction(5% = app. 200 words)
    - *Describe the context of your topic of interest, and your motivation to pursue said topic.*
2. Problem Statement (5% = app. 200 words)
    - *Define the problem*
      2.1 Research Question
    - *Once the problem is defined, what question would give you an appropriate answer to said problem (part of SG2)*
3. Literature Review / State of the Art (20% = app. 800 words)
    - *Use academic sources/articles to explain what we already know about this topic. At least 5 cited articles.*
4. Methodology (30% = app. 1200 words)
      4.1 Descriptive statistics
    - *Data source(s), variable description (labels), describe the variables (traditionally descriptive statistics include the mean, median, and standard deviation)*
    - 4.2 Analysis plan
    - *As given in SG4.*
5. Findings and Discussion (20% = app. 800 words)
    - *Report your analysis' findings and discuss what you have learned from them*
6. Conclusion (10% = app. 400 words)
    - *Give concluding remarks about your topic and findings. State the limitations of your study and avenues for future research (if you had all the resources available, what else would you have done?)*
7. Bibliography
    - *MLA, APA, or Harvard are all acceptable citation formats.*
    - *-The bibliography is not included in the word count.*

Significant deviations from total word count or insufficient bibliographic material will incur grade penalties.

Reading Material

Many, but not all, of the starred readings on this list are available on Canvas in the literature section of this course. Some are not there for copyright reasons. Some are not there because there is a link you can follow in the list below. Most of the starred readings should be relatively easy for you to find either by one of those methods, on the internet, or in the University library.

Task 1 and Task 2:

*Cowan,R. (2004). Network models of innovation and knowledge http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.318.6928&rep=rep1&type=pdf First half. (up to page 10)

*Kirman, A. P. (1992). Whom or what does the representative individual represent? *The Journal of Economic Perspectives*, 6(2), 117-136.

Fagiolo, G. Slides on Economic Networks: Theory and Empirics http://www.lem.sssup.it/fagiolo/files/nets_1.pdf

Task 3:

Newman, M. (2010). *Networks: An Introduction*. Oxford University Press, Inc. First chapter.

Fagiolo, G. Slides on Economic Networks: Theory and Empirics http://www.lem.sssup.it/fagiolo/files/nets_1.pdf

*Cowan, R. (2019). "Cowan-NetworkPreliminariesWithR.pdf"

Available on Canvas in the CourseModules section.

Task 5:

Mowery, D. C., Oxley, J. E., & Silverman, B. S. (1996). Strategic alliances and interfirm knowledge transfer. *Strategic management journal*, 17(S2), 77-91.

*Mowery, D. C., Oxley, J. E., & Silverman, B. S. 1998. Techno- logical overlap and interfirm cooperation: Implications for the resource-based view of the firm. *Research Policy*, 27: 507–523. Particularly the discussion up to page 511.

* Cowan, R. (2004). Network models of innovation and knowledge diffusion (No. 016). http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.318.6928&rep=rep1&type=pdf (see tasks 5&6) Second Part.

*Dutta, B., & Jackson, M. O. (2003). On the formation of networks and groups. In *Networks and Groups* (pp. 1-15). Springer Berlin Heidelberg. https://web.stanford.edu/~jacksonm/netintro.pdf

Task 6:

Birke, D. and G.M.P. Swann Network effects, network structure and consumer interaction in mobile telecommunications in Europe and Asia, *Journal of Economic Behavior & Organization* Vol. 76, Issue 2, November 2010, Pages 153-167

*Easley and Kleinberg Chapter 3 "Strong and Weak Ties" https://www.cs.cornell.edu/home/kleinber/networks-book/networks-book-ch03.pdf especially pages 65-69.

*Granovetter, M. S. (1973). The strength of weak ties. *American journal of sociology*, 78(6), 1360-1380.

Van Alstyne, M., & Brynjolfsson, E. (1996, December). Electronic communities: Global village or cyberbalkans?. In *Proceedings of the International Conference on Information Systems* (pp. 80-98).

Van Alstyne, M., & Brynjolfsson, E. (1996). Could the Internet balkanize science? *Science*, 274(5292), 1479.

Stainback, K. (2008). Social contacts and race/ethnic job matching. *Social Forces*, 87(2), 857-886. Particularly up to page 863 and the Conclusion section.

*Coleman, J. S. (1988). "Social capital in the creation of human capital." *American Journal of Sociology*, 94, S95-S120.

* R. Burt "Structural holes versus network closer as social capital" in Lin, Cook, Burt, eds, *Social Capital: Theory and Research*, 2000.

Task 7:

*Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. Nature, 393(6684), 440-442.

*Easley and Kleinberg *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Chapter 20, up to page 616 https://www.cs.cornell.edu/home/kleinber/networks-book/networks-book-ch20.pdf

*Milgram, S. The Small World Problem, *Psychology Today*, vol 1, May 1967, pp61-67.

Watts, D.J *Small Worlds: The dynamics of networks between order and randomness*.

Task 8:

Barabási, A. L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439), 509-512.

*Barabasi, A.L. *Linked*. Chapter 6 and 7.

*De Solla Price "Networks of Scientific Papers", *Science*, New Series, Vol. 149, No. 3683 (Jul. 30, 1965), pp. 510-515

Task 9:

Angelucci, M., G. De Giorgi, M. Rangel, and I. Rasul (2010): "Family Networks and School Enrollment: Evidence from a Randomized Social Experiment", *Journal of Public Economics*, 94(3-4)

Bandiera, O., and I. Rasul (2006): "Social Networks and Technology Adoption in Northern Mozambique", *Economic Journal*, 116(514), pp. 862-902.

Eagle, N., M. Macy and R. Claxton (2010) "Network diversity and Economic development" *Science*, Vol 328, 2 May, pp 1029-1031.

Opp, K. D. and C. Gern (1993): "Dissident groups, personal networks, and spontaneous cooperation: the East German revolution of 1989," *American Sociological Review* 58, 659-680.

Price, D. J. de S. (1965): "Networks of scientific papers," *Science* 149, 510-515.

Redner, S. (1998): "How popular is your paper? An empirical study of the citation distribution," *European Physical Journal* B 4, 131-134.

Granovetter, Mark S. "The Strength of Weak Ties." *American Journal of Sociology*, vol. 78, no. 6, 1973, pp. 1360–80, http://www.jstor.org/stable/2776392. Accessed 6 Apr. 2022.

Birke, D. and G.M.P. Swann Network effects, network structure and consumer interaction in mobile telecommunications in Europe and Asia*, Journal of Economic Behavior & Organization* Vol 76, Issue 2, November 2010, Pages 153-167

Milgram, S. The Small World Problem, *Psychology Today*, vol 1, May 1967, pp61-67.

Appendices

Student Code of Conduct

How do we wish to interact with each other? Here you will find a highlight of the values of the School of Business and Economics (SBE), which inspire our community and act as a

framework for the personal conduct of employees and students:

01 Respect

First of all, we should be able to count on each other: among staff, among students, and in cooperation with each other. Respect, reliability, taking responsibility and self-reflection are the core concepts here. This includes, but is not limited to, the use of appropriate language in both written and spoken communication.

02 Commitment

We are part of a community of motivated students and staff. We expect students and staff to express their intellectual passion through ambition, initiative and commitment. Staff and students must encourage each other in this process.

03 Professionalism

Staff and students contribute to an inspiring work and academic climate from their various fields of expertise, knowledge and experience. We value each other's contribution in word and gesture. Students and employees have many rights as well as obligations and we acknowledge these and apply them appropriately and professionally at all times.

04 Inclusivity

Our School is dedicated to creating an inclusive environment for everyone, regardless of race, ethnicity, religion, colour, nationality, sexual orientation, gender etc. The inclusion of all students and staff members is fundamental to our School's diverse and international character. English is used as the lingua franca at all times within our School.

05 Integrity

We are committed to protecting and guaranteeing academic integrity. This means taking exams independently and honestly, indicating sources when writing a paper, essay or thesis and always being truthful when filling out forms and other documentation.

This Code of Conduct offers the possibility to hold someone accountable for undesirable behaviour. For more details, including complaint procedures and potential sanctions, please consult our full code of conduct on SBE's Student Intranet.

Fraud and Plagiarism

In order to protect the reputation of the degrees that you – as students – receive, instances of cheating or plagiarism are treated extremely seriously.

Fraud, including plagiarism, is understood as a student's act or failure to act that makes it partially or fully impossible to correctly assess his/her knowledge, insight and skills. Plagiarism is understood as the presentation of one's own or other people's ideas or words without adequate reference to the source. Any assignment is an individual piece of work, which means that plagiarism is strictly forbidden. Equally, the use of mobile phones, communication devices or any other information carrier (whether the phone or other device is turned on or off, used or not used, etc. is irrelevant) during an examination is also forbidden.

If the Board of Examiners concludes that anything has occurred in an examination that makes it partially or fully impossible to correctly assess his/her knowledge, insight and skills, they may impose a sanction in accordance with SBE's policy on fraud, including plagiarism.

More information can be found on MySBE Intranet:
https://intranet.maastrichtuniversity.nl/en/cheating-and-plagiarism

Partial results

If students fail a course and have to redo it in the next academic year, they have two options:

1. Register for the exam only (assuming the typical case in which a student has failed the written exam but has passed other parts of the overall examination). Other partial grades remain valid and the final grade will be calculated using the weights that were used in the academic year in which the student followed the course.

2. Register for education. If a student decides to register for tutorials, (s)he will get new grades for those grades that are assessed during the tutorial meetings (e.g. participation, presentation). This makes sure that students that register for tutorials can't passively 'consume education', but instead are required to actively participate in all educational activities

Resits

Make clear that according to EER, exclusion is only permitted if:

The exclusion of a resit opportunity for (a) component(s) of (the examination of) an educational unit is only permitted if the aforementioned component(s), because of its/their nature can only be assessed within the relevant course period (e.g. participation, software labs, presentation).

Comments, objections and appeal

Students can find detailed information on these procedures on the Intranet: https://intranet.maastrichtuniversity.nl/en/comments-complaints-and-appeal as well as the EER: https://intranet.maastrichtuniversity.nl/en/education-and-examination-regulations

1. Comments

General) Comments regarding written examinations with open and/or closed questions (Multiple Choice) or parts thereof have to be submitted to the course coordinator according to the rules as published on the Intranet via https://examcomments.sbe.maastrichtuniversity.nl/. Deadline is 24:00 of the fifth day after the exam.

2. Inspection and objections

Within ten (10) working days of the publication of your examination results in the Student Portal >> My Courses >> Show results, you will be able to have a look at your assessed work (exam inspection).

You can request a copy of your multiple choice answer sheet via Surfyourself (SYS) in order to inspect your examination.

It is your course coordinator's responsibility to announce when and how you can do this. Usually, the date(s) will be posted with the (digital) publication of examination results or on the course page in My Courses.

If you can demonstrate that you were not able to make a complaint within the specified period due to circumstances beyond your control, you will be given another ten (10) working days period in which to do so.

If you do not agree with your assessment, you can submit objections in writing during the inspection, according to the instructions communicated by the course coordinator.

3. Complaints and appeal

You can send a complaint (not being a comment or an objection as mentioned above) regarding the course to the Complaints Service Point CSP: https://www.maastrichtuniversity.nl/support/during-your-studies/complaints-service-point-csp-one-service-desk-objections-appeals-or-0

If you and the course coordinator cannot reach an agreement or you are not satisfied with a decision of an examiner or the Board of Examiners, you can lodge an appeal to the Board of Appeal for Examinations (College van Beroep voor de Examens, CBE), within six (6) weeks after you received the decision. You may only appeal if a decision has gone directly against your interests. More information about what you may lodge an appeal against with the Board of Appeal for Examinations can be found on the Intranet and at the CSP link above.