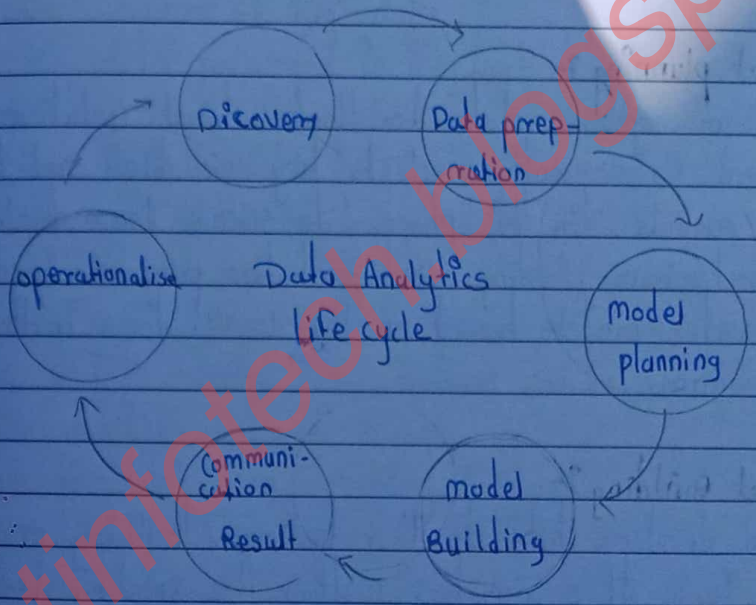
Assignment No :- 02

Lucky Page No.:
Date : / /

Q.1 Define data analytic life cycle.

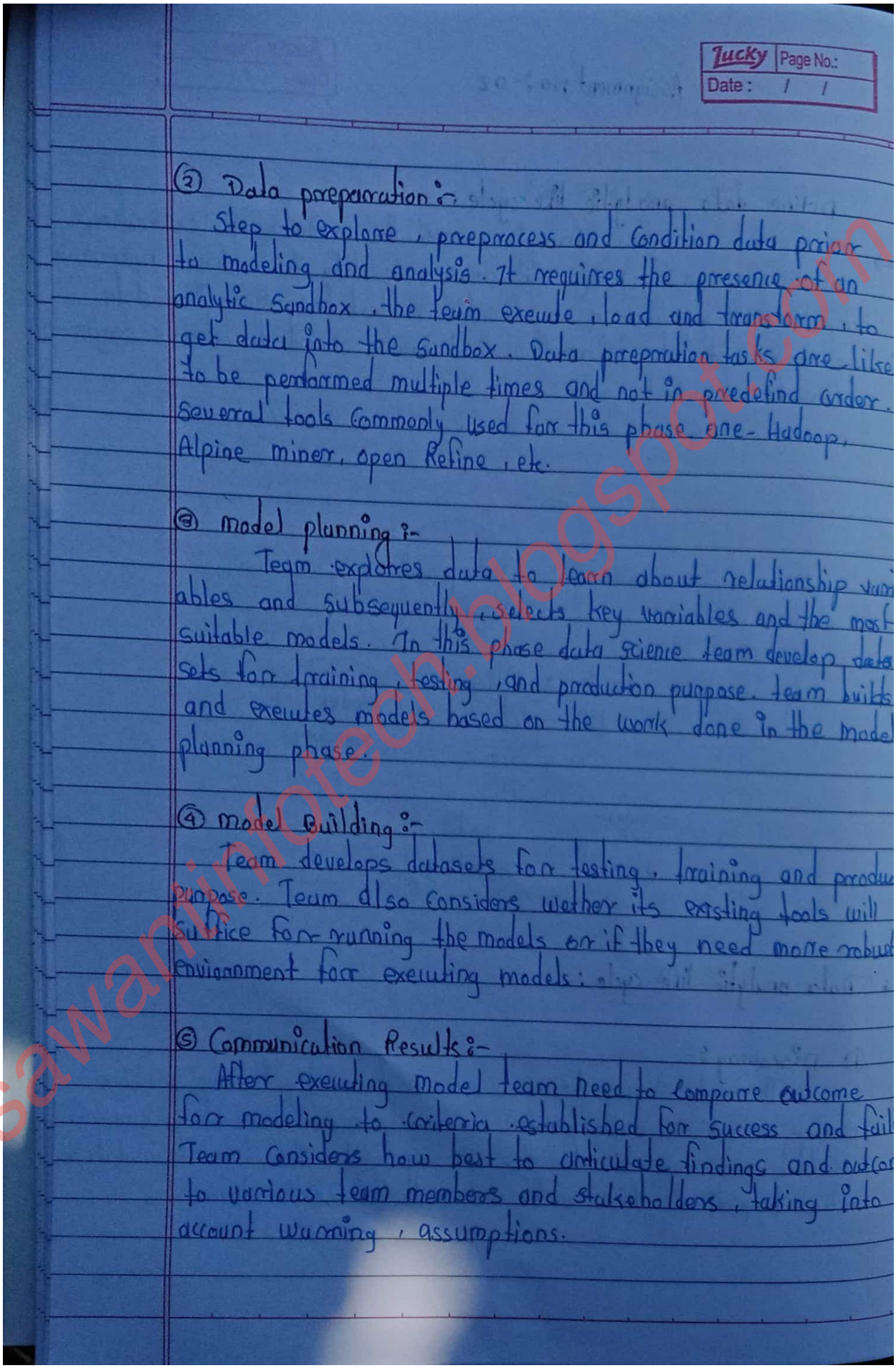
⟹ The data analytics life cycle covers the process of generating, collecting processing using and analyzing data to achieve corporate objectives. It provides a systematic method for managing data to convert it into information that can be used to achieve organizational and project goals.



* Data analytic life cycle :-

① Discovery :-

The data science team learn and investigate the problem develop context and understanding. Come to know about data sources needed and available for the the project. The team formulates initial hypothesis that can be later tested with data.



② Data preparation :-

Step to explore, preprocess and condition data prior to modeling and analysis. It requires the presence of an analytic sandbox, the team execute, load and transform, to get data into the sandbox. Data preparation tasks are likely to be performed multiple times and not in predefind order. Several tools commonly used for this phase are- Hadoop, Alpine miner, open Refine, etc.

③ model planning :-

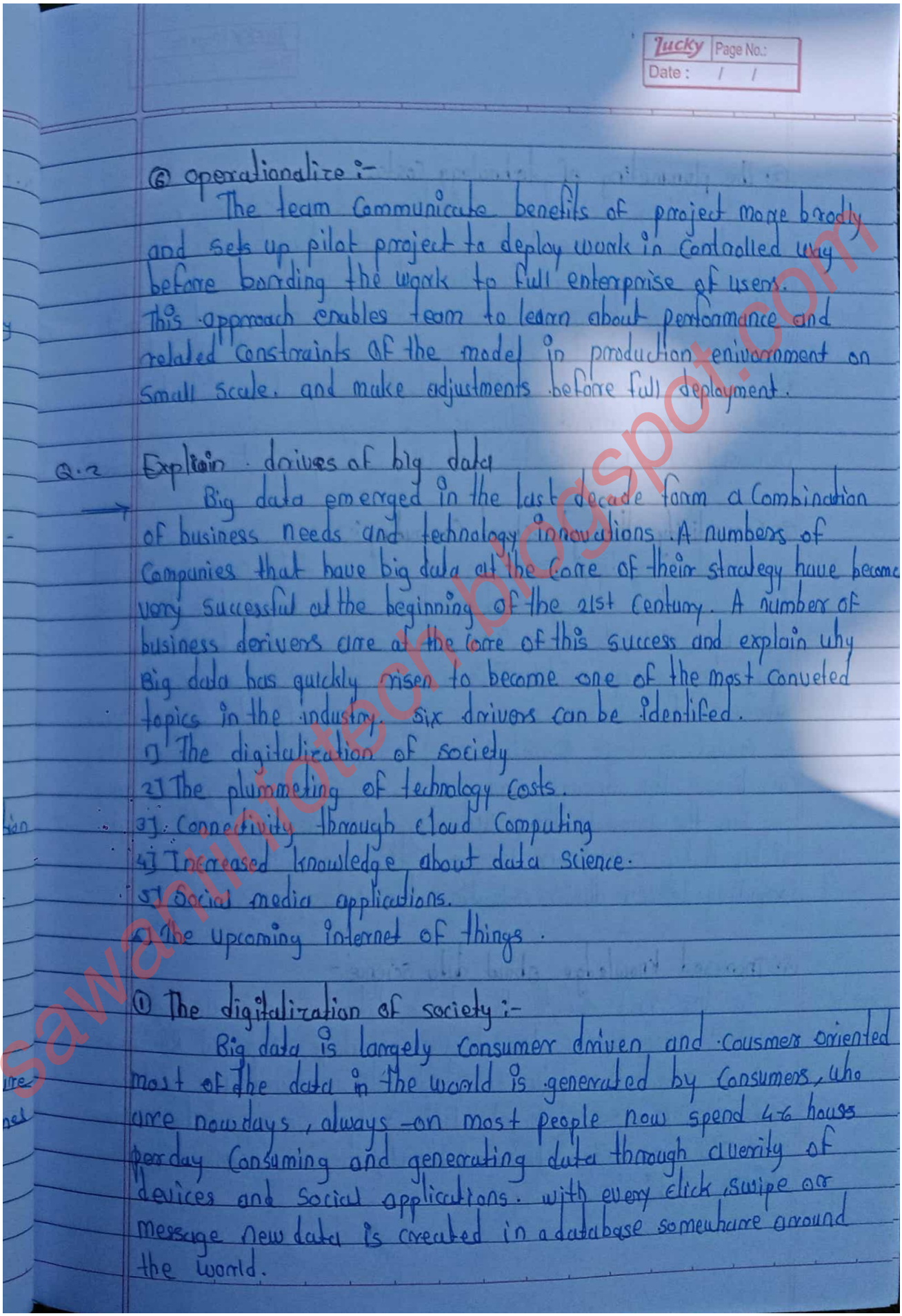
Team explores data to learn about relationship variables and subsequently selects key variables and the most suitable models. In this phase data science team develop data sets for training, testing, and production purpose. team builds and executes models based on the work done in the model planning phase.

④ model Building :-

Team develops datasets for testing, training and production purpose. Team also considers wether its existing tools will suffice for running the models or if they need more robust environment for executing models.

⑤ Communication Results :-

After executing model team need to compare outcome for modeling to criteria established for success and failure. Team considers how best to articulate findings and outcomes to various team members and stakeholders, taking into account warning, assumptions.



⑥ operationalize :-

The team communicate benefits of project more broadly and sets up pilot project to deploy work in controlled way before bording the work to full enterprise of users. This approach enables team to learn about performance and related constraints of the model in production environment on small scale. and make adjustments before full deployment.

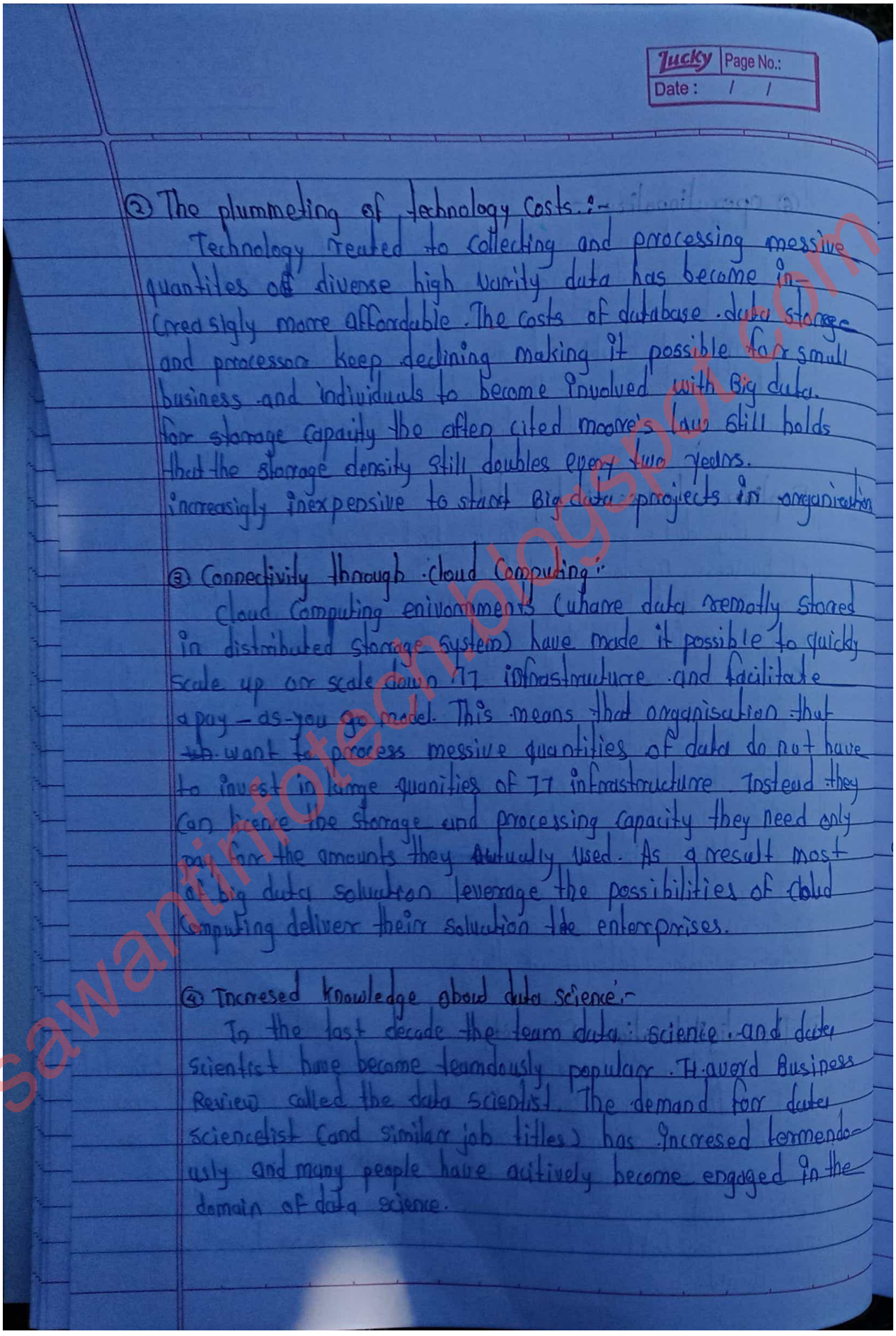
Q.2 Explain drives of big data

→ Big data emerged in the last decade form a combination of business needs and technology innovations. A numbers of companies that have big data at the core of their strategy have become very successful at the beginning of the 21st century. A number of business derivers are at the core of this success and explain why Big data has quickly risen to become one of the most conveted topics in the industry. Six drivers can be identified.

1] The digitalization of society
2] The plummeting of technology costs.
3] Connectivity through cloud computing
4] Increased knowledge about data science.
5] Social media applications.
6] The upcoming internet of things.

① The digitalization of society :-

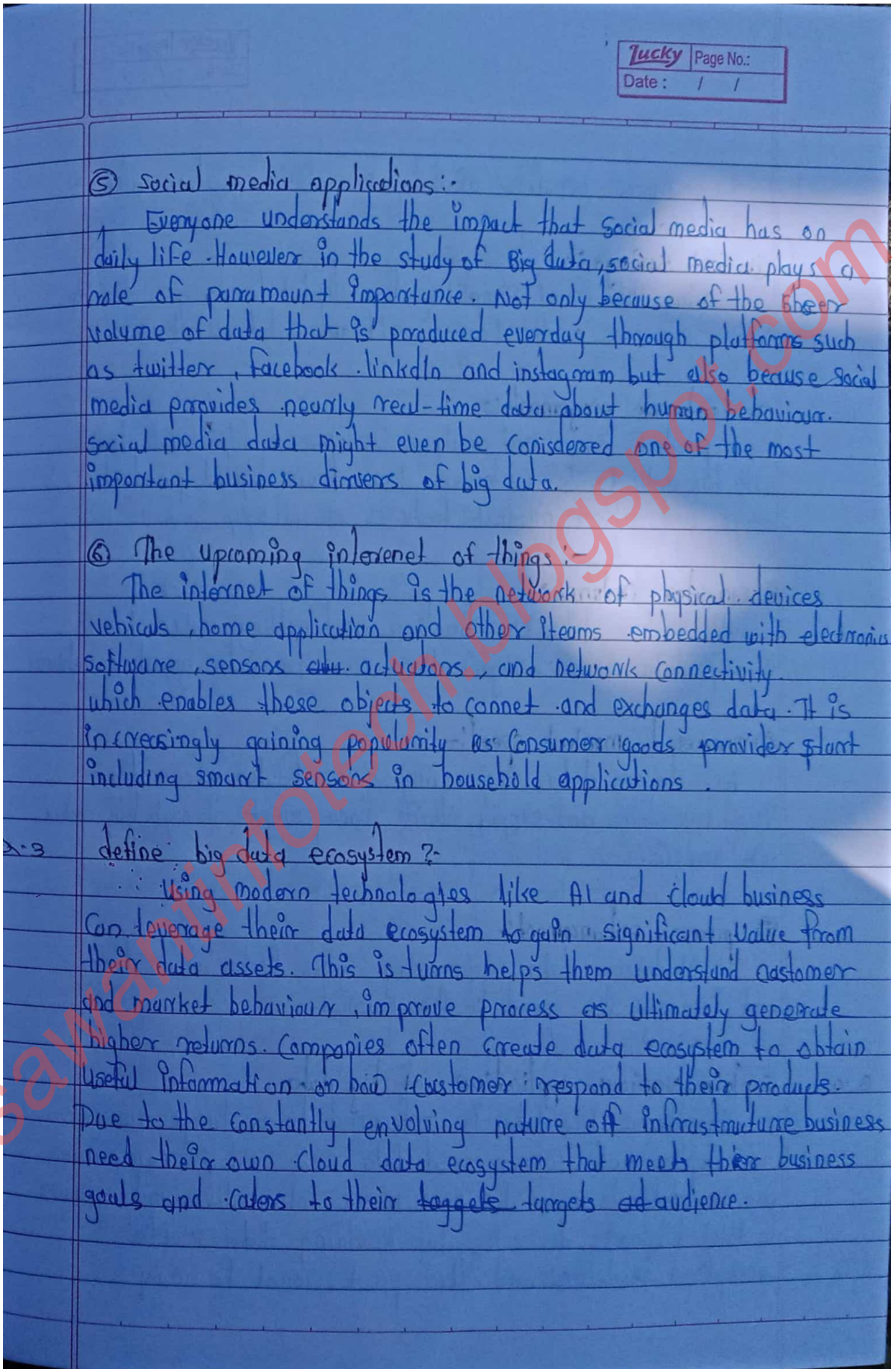
Big data is largely consumer driven and cousmer oriented most of the data in the world is generated by consumers, who are nowdays, always-on most people now spend 4.6 hours per day consuming and generating data through ubiquity of devices and social applications. with every click, swipe or message new data is created in a database somewhere around the world.



② The plummeting of technology costs :-
Technology related to collecting and processing massive quantites of diverse high variety data has become increasingly more affordable. The costs of database, data storage and processor keep declining making it possible for small business and individuals to become involved with Big data. For storage capacity the often cited moore's law still holds that the storage density still doubles every two years. increasingly inexpensive to start Big data projects in organisation

③ Connectivity through cloud Computing :-
Cloud Computing environments (where data remotely stored in distributed storage system) have made it possible to quickly scale up or scale down IT infrastructure and facilitate a pay-as-you-go model. This means that organisation that want to process massive quantities of data do not have to invest in large quantities of IT infrastructure. Instead they can licence the storage and processing capacity they need only pay for the amounts they actually used. As a result most of big data solution leverage the possibilities of cloud computing deliver their solution the enterprises.

④ Increased knowledge about data science :-
In the last decade the term data science and data scientist have become tremendously popular. Harvard Business Review called the data scientist. The demand for data scientist (and similar job titles) has increased tremendously and many people have actively become engaged in the domain of data science.



⑤ Social media applications :-

Everyone understands the impact that social media has on daily life. However in the study of Big Data, social media plays a role of paramount importance. Not only because of the sheer volume of data that is produced everyday through platforms such as twitter, facebook, linkdin and instagram but also because social media provides nearly real-time data about human behaviour. Social media data might even be considered one of the most important business dimsens of big data.

⑥ The upcoming internet of things :-

The internet of things is the network of physical devices vehicals, home application and other iteams embedded with electronics software, sensors actuators, and network connectivity which enables these objects to connet and exchanges data. It is increasingly gaining popularity as consumer goods provider start including smart sensors in household applications.

Q.3 define big data ecosystem ?-

Using modern technologies like AI and cloud business can leverage their data ecosystem to gain significant value from their data assets. This is turns helps them understand customer and market behaviour, improve process as ultimately generate higher returns. Companies often create data ecosystem to obtain useful information on how customer respond to their products. Due to the constantly envolving nature of infrastructure business need their own cloud data ecosystem that meets their business goals and caters to their ~~toggets~~ targets of audience.

Some more benefits of data ecosystem creation are:-

① Cost savings :-
Using the cloud simplifies the digital landscape due to fewer expenses on the transition from a data warehouse.

② Customer engagements :-
Product teams can analyze customer's like and dislike with the data trails customers leave from digital product usage and alter product features accordingly.

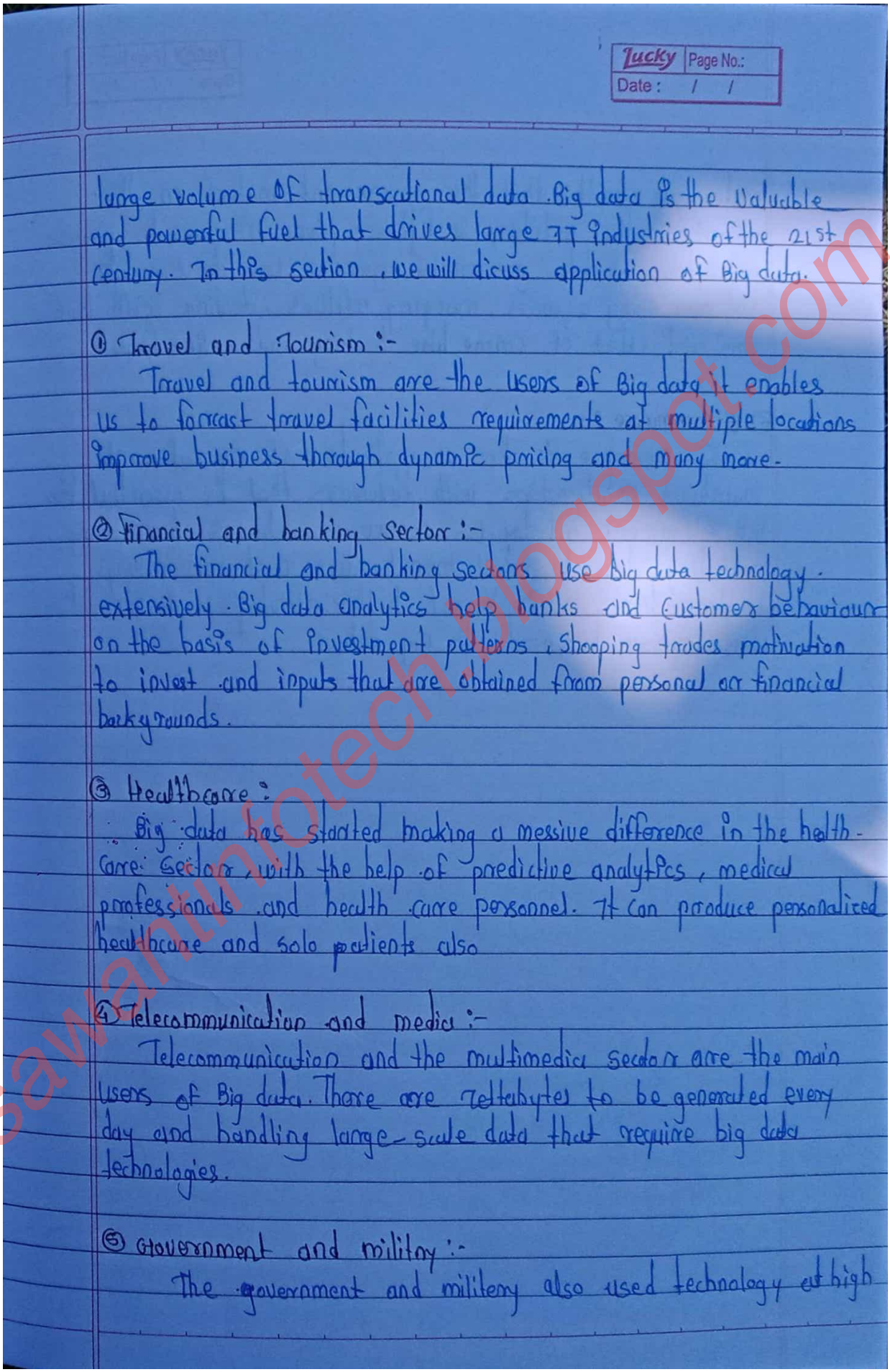
③ Greater returns :-
With higher data monetization and getting value from legacy data stores, organizations can generate greater returns.

④ Increased market speed :-
AI-driven data engineering gives faster information like inventory and supply chain management through big data set analysis.

⑤ process enhancement :-
AI driven data engineering gives faster information and increase its speed to the market.

Q.4 Explain various application of big data.

The term Big data is referred to as large amount of complex and unprocessed data. Now a day's companies use Big data to make business more informative and allows to take business decisions by enabling data scientists, analytical modelers and other professional to analyse



large volume of transactional data. Big data is the valuable and powerful fuel that drives large IT industries of the 21st century. In this section, we will dicuss application of Big data.

① Travel and Tourism :-

Travel and tourism are the users of Big data it enables us to forcast travel facilities requirements at multiple locations improve business through dynamic pricing and many more.

② Financial and banking sector :-

The financial and banking sectors use Big data technology extensively. Big data analytics help banks and customer behaviour on the basis of investment patterns, shooping trades motivation to invest and inputs that are obtained from personal or financial backgrounds.

③ Healthcare :

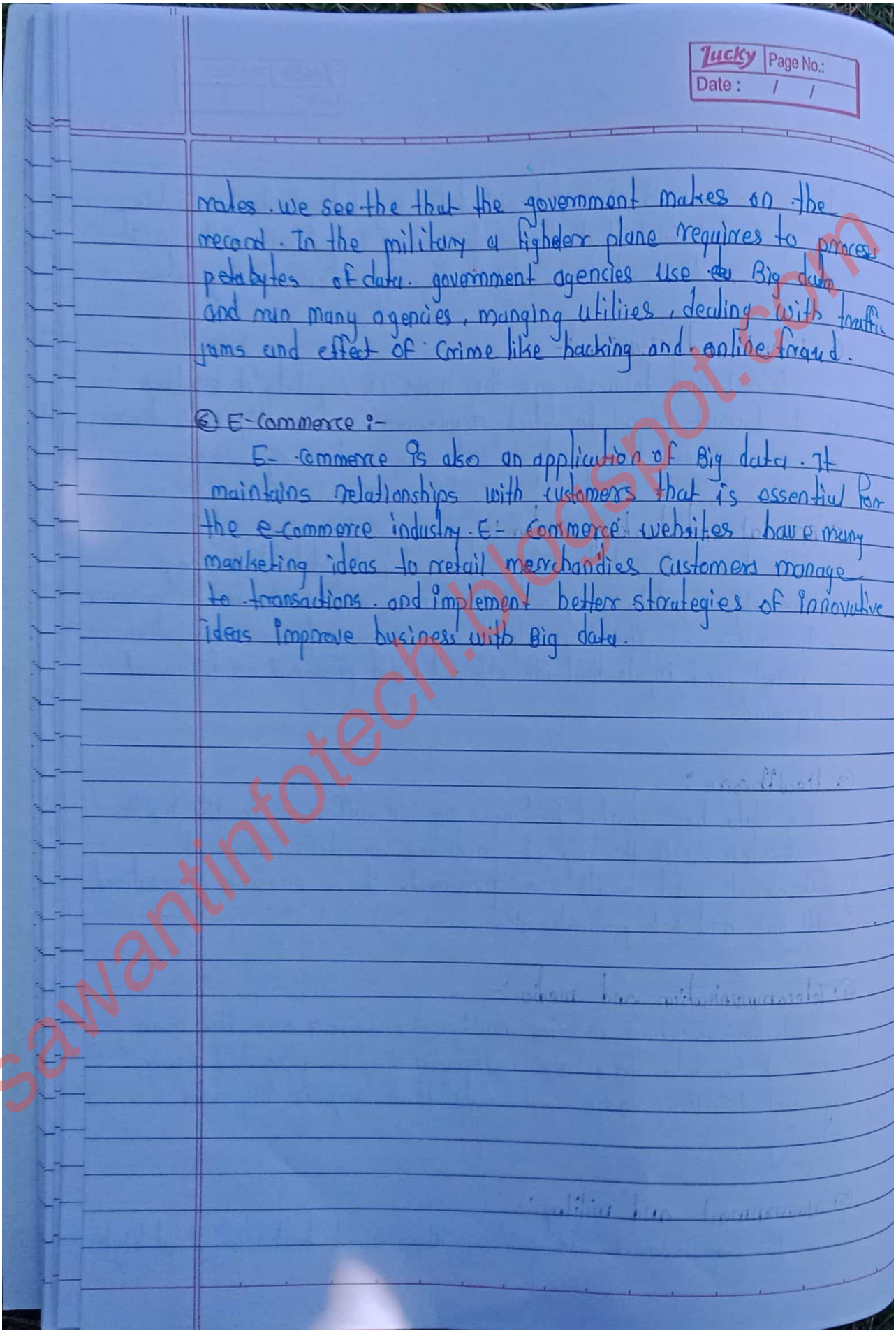
Big data has started making a messive difference in the health-care sector with the help of predictive analytics, medical professionals and health care personnel. It can produce personalized healthcare and solo patients also

④ Telecommunication and media :-

Telecommunication and the multimedia secdor are the main users of Big data. There are zettabytes to be generated every day and handling large-scale data that require big data technologies.

⑤ Government and military :-

The government and militery also used technology at high



Lucky Page No.:
Date : / /

makes. we see the that the government makes on the record. In the military a fighter plane requires to process petabytes of data. government agencies use Big data and run many agencies, manging utilities, dealing with traffic jams and effect of crime like hacking and online fraud.

② E-Commerce :-

E- commerce is also an application of Big data. It maintains relationships with customers that is essential for the e-commerce industry. E- commerce websites have many marketing ideas to retail merchandies customers manage to transactions. and implement better strategies of innovative ideas Improve business with Big data.