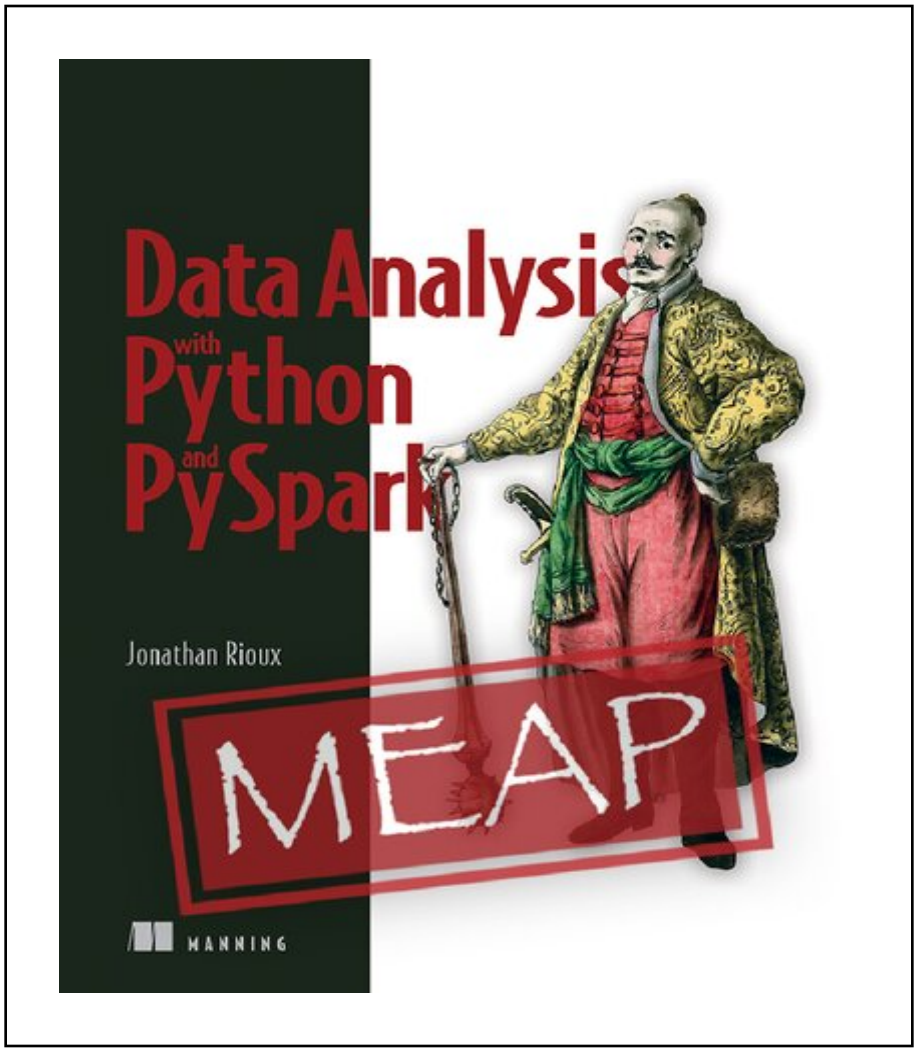
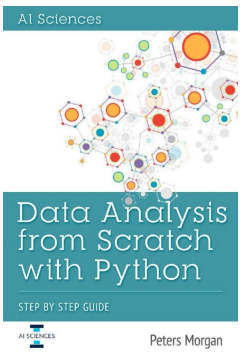
Data Analysis with Python and PySpark (MEAP V07)
Jonathan Rioux

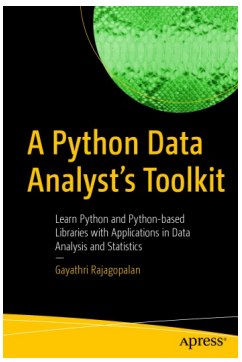


More products digital (pdf, epub, mobi) instant download maybe you interests ...



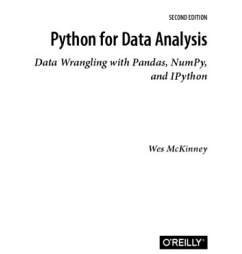
Data Analysis from Scratch with Python Peters Morgan

https://textbookfull.com/product/data-analysis-from-scratch-with-python-peters-morgan/



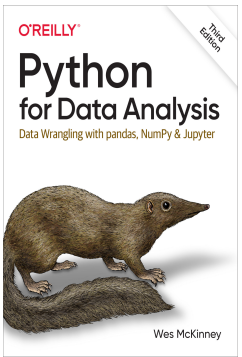
A Python Data Analyst's Toolkit: Learn Python and Python-based Libraries with Applications in Data Analysis and Statistics Gayathri Rajagopalan

https://textbookfull.com/product/a-python-data-analysts-toolkit-learn-python-and-python-based-libraries-with-applications-in-data-analysis-and-statistics-gayathri-rajagopalan/



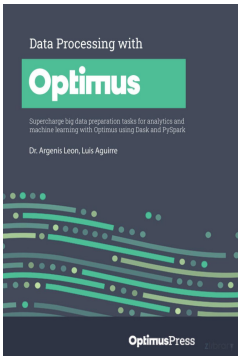
Python for Data Analysis Data Wrangling with Pandas NumPy and IPython Wes Mckinney

https://textbookfull.com/product/python-for-data-analysis-data-wrangling-with-pandas-numpy-and-ipython-wes-mckinney/



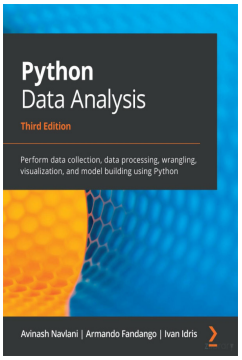
Python for Data Analysis Data Wrangling with pandas NumPy and Jupyter 3rd Edition Wes Mckinney

https://textbookfull.com/product/python-for-data-analysis-data-wrangling-with-pandas-numpy-and-jupyter-3rd-edition-wes-mckinney/



Data Processing with Optimus: Supercharge big data preparation tasks for analytics and machine learning with Optimus using Dask and PySpark Leon

https://textbookfull.com/product/data-processing-with-optimus-supercharge-big-data-preparation-tasks-for-analytics-and-machine-learning-with-optimus-using-dask-and-pyspark-leon/

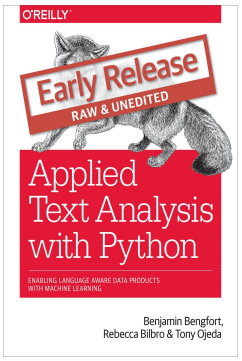


Python Data Analysis: Perform data collection, data processing, wrangling, visualization, and model building using Python 3rd Edition Avinash Navlani

https://textbookfull.com/product/python-data-analysis-perform-data-collection-data-processing-wrangling-visualization-and-model-building-using-python-3rd-edition-avinash-navlani/

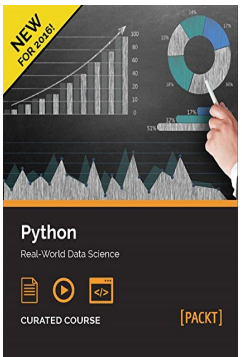
Learning Data Mining with Python Layton

https://textbookfull.com/product/learning-data-mining-with-python-layton/



Applied Text Analysis with Python Enabling Language Aware Data Products with Machine Learning 1st Edition Benjamin Bengfort

https://textbookfull.com/product/applied-text-analysis-with-python-enabling-language-aware-data-products-with-machine-learning-1st-edition-benjamin-bengfort/



Data Science with Python 1st Edition Coll.

https://textbookfull.com/product/data-science-with-python-1st-edition-coll/

Data Analysis with Python and PySpark

Jonathan Rioux

MEAP

MANNING

MEAP Edition
Manning Early Access Program
Data Analysis with Python and PySpark
Version 7

For more information on this and other Manning titles go to
manning.com

welcome

Thank you for purchasing the MEAP for *Data Analysis with Python and PySpark.* It is a lot of fun (and work!) and I hope you'll enjoy reading it as much as I am enjoying writing the book.

My journey with PySpark is pretty typical: the company I used to work for migrated their data infrastructure to a data lake and realized along the way that their usual warehouse-type jobs didn't work so well anymore. I spent most of my first months there figuring out how to make PySpark work for my colleagues and myself, starting from zero. This book is very influenced by the questions I got from my colleagues and students (and sometimes myself). I've found that combining practical experience through real examples with a little bit of theory brings not only proficiency in using PySpark, but also how to build better data programs. This book walks the line between the two by explaining important theoretical concepts without being too laborious.

This book covers a wide range of subjects, since PySpark is itself a very versatile platform. I divided the book into three parts.

- Part 1: Walk teaches how PySpark works and how to get started and perform basic data manipulation.
- Part 2: Jog builds on the material contained in Part 1 and goes through more advanced subjects. It covers more exotic operations and data formats and explains more what goes on under the hood.
- Part 3: Run tackles the *cooler* stuff: building machine learning models at scale, squeezing more performance out of your cluster, and adding functionality to PySpark.

To have the best time possible with the book, you should be at least comfortable using Python. It isn't enough to have learned another language and transfer your knowledge into Python. I cover more niche corners of the language when appropriate, but you'll need to do some research on your own if you are new to Python.

Furthermore, this book covers how PySpark can interact with other data manipulation frameworks (such as Pandas), and those specific sections assume basic knowledge of Pandas.

Finally, for some subjects in Part 3, such as machine learning, having prior exposure will help you breeze through. It's hard to strike a balance between "not enough explanation" and "too much explanation"; I do my best to make the right choices.

Your feedback is key in making this book its best version possible. I welcome your comments and thoughts in the liveBook discussion forum.

Thank you again for your interest and in purchasing the MEAP!

—Jonathan Rioux

brief contents

1 Introduction

In this chapter, you will learn

- What is PySpark
- Why PySpark is a useful tool for analytics
- The versatility of the Spark platform and its limitations
- PySpark's way of processing data

According to pretty much every news outlet, data is everything, everywhere. It's the new oil, the new electricity, the new gold, plutonium, even bacon! We call it powerful, intangible, precious, dangerous. I prefer calling it *useful in capable hands*. After all, for a computer, any piece of data is a collection of zeroes and ones, and it is our responsibility, as users, to make sense of how it translates to something useful.

Just like oil, electricity, gold, plutonium and bacon (especially bacon!), our appetite for data is growing. So much, in fact, that computers aren't following. Data is growing in size and complexity, yet consumer hardware has been stalling a little. RAM is hovering for most laptops at around 8 to 16 GB, and SSD are getting prohibitively expensive past a few terabytes. Is the solution for the burgeoning data analyst to triple-mortgage his life to afford top of the line hardware to tackle Big Data problems?

Introducing Spark, and its companion PySpark, the unsung heroes of large-scale analytical workloads. They take a few pages of the supercomputer playbook— powerful, but manageable, compute units meshed in a network of machines— and bring it to the masses. Add on top a powerful set of data structures ready for any work you're willing to throw at them, and you have a tool that will *grow* (pun intended) with you.

This book is great introduction to data manipulation and analysis using PySpark. It tries to cover

just enough theory to get you comfortable, while giving enough opportunities to practice. Each chapter except this one contains a few exercices to anchor what you just learned. The exercises are all solved and explained in Appendix A.

1.1 What is PySpark?

What's in a name? Actually, quite a lot. Just by separating PySpark in two, one can already deduce that this will be related to Spark and Python. And it would be right!

At the core, PySpark can be summarized as being the Python API to Spark. While this is an accurate definition, it doesn't give much unless you know the meaning of Python and Spark. If we were in a video game, I certainly wouldn't win any prize for being the most useful NPC. Let's continue our quest to understand what is PySpark by first answering *What is Spark?*.

1.1.1 You saw it coming: What is Spark?

Spark, according to their authors, is a *unified analytics engine for large-scale data processing*. This is a very accurate definition, if a little dry.

Digging a little deeper, we can compare Spark to an *analytics factory*. The raw material— here, data— comes in, and data, insights, visualizations, models, you name it! comes out.

Just like a factory will often gain more capacity by increasing its footprint, Spark can process an increasingly vast amount of data by *scaling out* instead of *scaling up*. This means that, instead of buying thousand of dollars of RAM to accommodate your data set, you'll rely instead of multiple computers, splitting the job between them. In a world where two modest computers are less costly than one large one, it means that scaling out is less expensive than up, keeping more money in your pockets.

The problem with computers is that they crash or behave unpredictably once in a while. If instead of one, you have a hundred, the chance that at least one of them go down is now much higher.[1] Spark goes therefore through a lot of hoops to manage, scale, and babysit those poor little sometimes unstable computers so you can focus on what you want, which is to work with data.

This is, in fact, one of the weird thing about Spark: it's a good tool because of what you can do with it, but especially because of what you *don't have to do* with it. Spark provides a powerful API[2] that makes it look like you're working with a cohesive, non-distributed source of data, while working hard in the background to optimize your program to use all the power available. You therefore don't have to be an expert at the arcane art of distributed computing: you just need to be familiar with the language you'll use to build your program. This leads us to…

1.1.2 PySpark = Spark + Python

PySpark provides an entry point to Python in the computational model of Spark. Spark itself is coded in Scala, a language very powerful if a little hard to grasp. In order to meet users where they are, Spark also provides an API in Java, Python and R. The authors did a great job at providing a coherent interface between language while preserving the idiosyncrasies of the language where appropriate. Your PySpark program will therefore by quite easy to read by a Scala/Spark programmer, but also to a fellow Python programmer who hasn't jumped into the deep end (yet).

Python is a dynamic, general purpose language, available on many platforms and for a variety of tasks. Its versatility and expressiveness makes it an especially good fit for PySpark. The language is one of the most popular for a variety of domains, and currently is a major force in data analysis and science. The syntax is easy to learn and read, and the amount of library available means that you'll often find one (or more!) who's just the right fit for your problem.

1.1.3 Why PySpark?

There are no shortage of libraries and framework to work with data. Why should one spend their time learning PySpark specifically?

PySpark packs a lot of advantages for modern data workloads. It sits at the intersection of fast, expressive and versatile. Let's explore those three themes one by one.

PYSPARK IS FAST

If you search "Big Data" in a search engine, there is a very good chance that Hadoop will come within the first few results. There is a very good reason to this: Hadoop popularized the famous *MapReduce* framework that Google pioneered in 2004 and is now a staple in Data Lakes and Big Data Warehouses everywhere.

Spark was created a few years later, sitting on Hadoop's incredible legacy. With an aggressive query optimizer, a judicious usage of RAM and some other improvements we'll touch on in the next chapters, Spark can run up to 100x faster than plain Hadoop. Because of the integration between the two frameworks, you can easily switch your Hadoop workflow to Spark and gain the performance boost without changing your hardware.

PYSPARK IS EXPRESSIVE

Beyond the choice of the Python language, one of the most popular and easy-to-learn language, PySpark's API has been designed from the ground up to be easy to understand. Most programs read as a descriptive list of the transformations you need to apply to the data, which makes them easy to reason about. For those familiar with functional programming languages, PySpark code is conceptually closer to the "pipe" abstraction rather than pandas, the most popular in-memory DataFrame library.

You will obviously see many examples through this book. As I was writing those examples, I was pleased about how close to my initial (pen and paper) reasoning the code ended up looking. After understanding the fundamentals of the framework, I'm confident you'll be in the same situation.

PYSPARK IS VERSATILE

There are two components to this versatility. First, there is the *availability* of the framework. Second, there is the diversified *ecosystem* surrounding Spark.

PySpark is everywhere. All three major cloud providers have a managed Hadoop/Spark cluster as part of their offering, which means you have a fully provisioned cluster at a click of a few buttons. You can also easily install Spark on your own computer to nail down your program before scaling on a more powerful cluster. Appendix B covers how to get your own local Spark running, while Appendix C will walk through the current main cloud offerings.

PySpark is open-source. Unlike some other analytical software, you aren't tied to a single company. You can inspect the source code if you're curious, and even contribute if you have an idea for a new functionality or find a bug. It also gives a low barrier to adoption: download, learn, profit!

Finally, Spark's eco-system doesn't stop at PySpark. There is also an API for Scala, Java, R, as well as a state-of-the-art SQL layer. This makes it easy to write a polyglot program in Spark. A Java software engineer can tackle the ETL pipeline in Spark using Java, while a data scientist can build a model using PySpark.

WHERE PYSPARK FELL SHORT

It would be awesome if PySpark was The Answer to every data problem. Unfortunately, there are some caveats. None of them are a deal-breakers, but they are to be considered when you're selecting a framework for your next project.

PySpark isn't the right choice if you're dealing with small data sets. Managing a distributed cluster comes with some overhead, and if you're just using a single node, you're paying the price but aren't using the benefits. As an example, a PySpark shell will take a few seconds to launch:

this is often more than enough time to process data that fits within your RAM.

PySpark also has a disadvantage when it comes to the Java and Scala API. Since Spark is at the core a Scala program, Python code have to be translated to and from JVM[3] instructions. While more recent versions have been bridging that gap pretty well, pure Python translation, which happens mainly when you're defining your own User Defined Functions (UDF), will perform slower. We will cover UDF and some ways to mitigate the performance problem in Chapter 8.

Finally, while programming PySpark can feel easy and straightforward, managing a cluster can be a little arcane. Spark is a pretty complicated piece of software, and while the code base matured remarkably over the past few years, the days where scaling a 100-machine cluster and manage it as easily as a single node are far ahead. We will cover some of the developer-facing configuration and problems in the Chapter about performance, but for hairier problems, do what I do: befriend your dev ops.

1.1.4 Your very own factory: how PySpark works

In this section, I will explain how Spark processes a program. It can be a little odd to present the workings and underpinnings of a system that we claimed, a few paragraphs ago, hides that complexity. We still think that having a working knowledge of how Spark is set up, how it manages data and how it optimizes queries is very important. With this, you will be able to reason with the system, improve your code and figure out quicker when it doesn't perform the way you want.

If we're keeping the factory analogy, we can imagine that the cluster of computer where Spark is sitting on is the building.

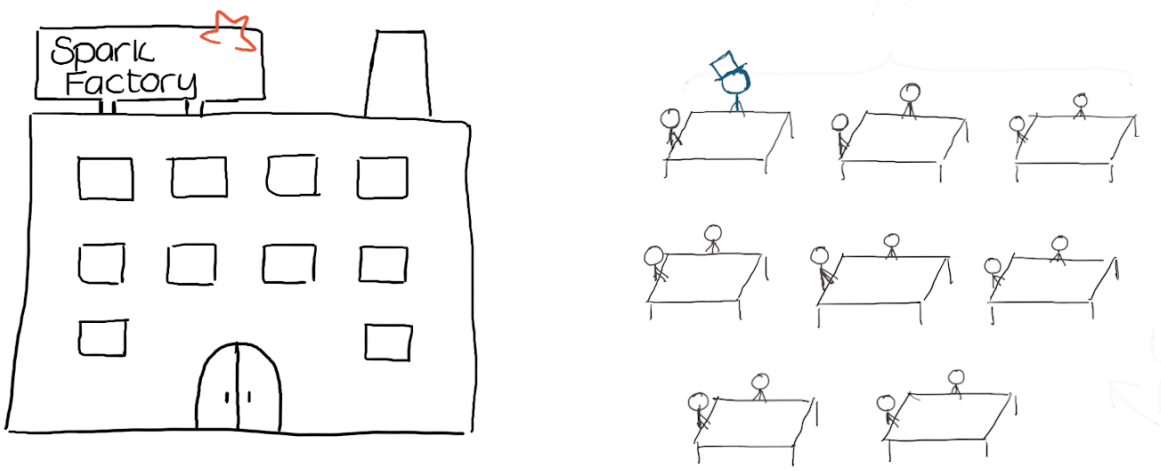


Figure 1.1 A totally relatable data factory, outside and in.

If we look at , we can see two different way to interpret a data factory. On the left, we see how it looks like from the outside: a cohesive unit where projects come in and results comes out. This is what it will appear to you most of the time. Under the hood, it looks more like on the right: you have some workbenches where some workers are assigned to. The workbenches are like the computers in our Spark cluster: there is a fixed amount of them, and adding or removing some is easy but needs to be planned in advance. The workers are called *executors* in Spark's literature: they are the one performing the actual work on the machines.

One of the little workers looks spiffier than the other. That top hat definitely makes him stand out of the crowd. In our data factory, he's the manager of the work floor. In Spark terms, we call this the *master*. In the spirit of the open work-space, he shares one of the workbenches with his fellow employees. The role of the master is crucial to the efficient execution of your program, so is dedicated to this.

1.1.5 Some physical planning with the cluster manager

Upon reception of the task, which is called a *driver program* in the Spark world, the factory starts running. This doesn't mean that we get straight to processing! Before that, the cluster need to *plan the capacity* it will allocate for your program. The entity, or program, taking care of this is aptly called the *cluster manager*. In our factory, this cluster manager will look at the workbenches with available space and secure as many as necessary, then start hiring workers to fill the capacity. In Spark, it will look at the machines with available computing resources and secure what's necessary, before launching the required number of executors across them.

> **NOTE** Spark provides its own cluster manager, but can also play well with other ones when working in conjunction with Hadoop or another Big Data platform. We will definitely discuss the intricacies of managing the cluster manager (pun intended) in the chapter about performance, but in the meantime, if you read about YARN or Mesos in the wild, know that they are two of the most popular nowadays.

Any directions about capacity (machines and executors) are encoded in a `SparkContext` object which represents the connection to our Spark cluster. Since our instructions didn't mention any specific capacity, the cluster manager will allocate the default capacity prescribed by our Spark installation.

We're off to a great start! We have a task to accomplish, and the capacity to accomplish it. What's next? Let's get working!

1.1.6 A factory made efficient through a lazy manager

Just like in a large-scale factory, you don't go to each employee and give them a list of tasks. No, here, you'll *provide your list of steps to the manager* and let them deal with it. In Spark, the manager/*master* takes your instructions (carefully written in Python code), translate them in Spark steps and then process them across the worker. The master will also manage which *worker* (more on them in a bits) has which slice of the data, and make sure that you don't lose some bits in the process.

Your manager/master has all the qualities a good manager has: smart, cautious and lazy. Wait, what? You read me right. *Laziness* in a programming context— and one could argue in the real world too— can actually be very good thing. Every instruction you're providing in Spark can be classified in two categories: transformations and actions. *Actions* are what many programming languages would consider IO. Actions includes, but are not limited to:

- Printing information on the screen
- Writing data to a hard drive or cloud bucket

In Spark, we'll see those instructions most often via the `show` and `write` methods, as well as other calling those two in their body.

Transformations are pretty much everything else. Some examples of transformation are:

- Adding a column to a table
- Performing an aggregation according to certain keys
- Computing summary statistics on a data set
- Training a Machine Learning model on some data

Why the distinction, you might ask? When thinking about computation over data, you, as the developer, are only concerned about the computation leading to an action. You'll always interact with the results of an action, because this is something you can see. Spark, with his lazy computation model, will take this to the extreme and will avoid performing data work until an action triggers the computation chain. Before that, the master will store (or *cache*) your instructions. This way of dealing with computation has many benefits when dealing with large scale data.

First, storing instructions in memory takes much less space than storing intermediate data results. If you are performing many operations on a data set and are materializing the data each step of the way, you'll blow your storage much faster although you don't need the intermediate results. We can all argue that less waste is better.

Second, by having the full list of tasks to be performed available, the master can optimize the work between executors much more efficiently. It can use information available at run-time, such as the node where specific parts of the data are located. It can also re-order and eliminate useless

transformations if necessary.

Finally, during interactive development, you don't have to submit a huge block of commands and wait for the computation to happen. Instead, you can iteratively build your chain of transformation, one at the time, and when you're ready to launch the computation (like during your coffee break), you can add an action and let Spark work its magic.

Lazy computation is a fundamental aspect of Spark's operating model and part of the reason it's so fast. Most programming languages, including Python, R and Java, are eagerly evaluated. This means that they process instructions as soon as they receive them. If you have never worked with a lazy language before, it can look a little foreign and intimidating. If this is the case, don't worry: we'll weave practical explanations and implications of that laziness during the code examples when relevant. You'll be a lazy pro in no time!

NOTE Reading data, although clearly being I/O, is considered a transformation by Spark. This is due to the fact that reading data doesn't perform any visible work to the user. You therefore won't read data until you need to display or write it somewhere.

What's a manager without competent employees? Once the task, with its action, has been received, the master starts allocating data to what Spark calls *executors*. Executors are processes that run computations and store data for the application. Those executors sit on what's called a *worker node*, which is the actual computer. In our factory analogy, an executor would be an employee performing the work, while the worker node would be a workbench where many employees/executors can work. If we recall , our master wears a top hat and sits with his employees/workers at one of the workbenches.

That concludes our factory tour. Let's summarize our typical PySpark program.

We first encode our instructions in Python code, forming a driver program.

When submitting our program (or launching a PySpark shell), the cluster manager allocates resources for us to use. Those will stay constant for the duration of the program.

The master ingests your code and translate it into Spark instructions. Those instructions are either transformations or actions.

Once the master reaches an action, it optimizes the whole computation chain and splits the work between executors. Executors are processes performing the actual data work and they reside on machines labelled worked nodes.

That's it! As we can see, the overall process is quite simple, but it's obvious that Spark hides a lot of the complexity arising from efficient distributed processing. For a developer, this means

shorter and clearer code, and a faster development cycle.

1.2 What will you learn in this book?

This book will use PySpark to solve a variety of tasks a data analyst, engineer or scientist will encounter during his day to day life. We will therefore

- read and write data from (and to) a variety of sources and formats;
- deal with messy data with PySpark's data manipulation functionality;
- discover new data sets and perform exploratory data analysis;
- build data pipelines that transform, summarize and get insights from data in an automated fashion;
- test, profile and improve your code;
- troubleshoot common PySpark errors, how to recover from them and avoid them in the first place.

After covering those fundamentals, we'll also tackle different tasks that aren't as frequent, but are interesting and an excellent way to showcase the power and versatility of PySpark.

- We'll perform Network Analysis using PySpark's own graph representation
- We'll build Machine Learning models, from simple throwaway experiments to Deep Learning goodness
- We'll extend PySpark's functionality using user defined functions, and learn how to work with other languages

We are trying to cater to many potential readers, but are focusing on people with little to no exposure to Spark and/or PySpark. More seasoned practitioners might find useful analogies for when they need to explain difficult concepts and maybe learn a thing or two!

The book focuses on Spark version 2.4, which is currently the most recent available. Users on older Spark versions will be able to go through most of the code in the book, but we definitely recommend using at least Spark 2.0+.

We're assuming some basic Python knowledge: some useful concepts are outlined in Appendix D. If you feel for a more in-depth introduction to Python, I recommend *The Quick Python Book*, by Naomi Ceder (Manning, 2018).

1.3 What do I need to get started?

In order to get started, the only thing absolutely necessary is a working installation of Spark. It can be either on your computer (Appendix B) or using a cloud provider (Appendix C). Most examples in the book are doable using a local installation of Spark, but some will require more horsepower and will be identified as such.

A code editor will also be very useful for writing, reading and editing scripts as you go through

the examples and craft your own programs. A Python-aware editor, such as PyCharm, is a nice-to-have but is in no way necessary. Just make sure it saves your code without any formatting: don't use Microsoft Word to write your programs!

The book's code examples are available on GitHub, so Git will be a useful piece of software to have. If you don't know git, or don't have it handy, GitHub provides a way to download all the book's code in a Zip file. Make sure you check regularly for updates!

Finally, I recommend that you have an analog way of drafting your code and schema. I am a compulsive note-taker and doodler, and even if my drawing are very basic and crude, I find that working through a new piece of software via drawings helps in clarifying my thoughts. This means less code re-writing, and a happier programmer! Nothing spiffy, some scrap paper and a pencil will do wonders.

1.4 Summary

- PySpark is the Python API for Spark, a distributed framework for large-scale data analysis. It provides the expressiveness and dynamism of the Python programming language to Spark.
- PySpark provides a full-stack analytics workbench. It has an API for data manipulation, graph analysis, streaming data as well as machine learning.
- Spark is fast: it owes its speed to a judicious usage of the RAM available and an aggressive and lazy query optimizer.
- Spark provides bindings for Python, Scala, Java, and R. You can also use SQL for data manipulation.
- Spark uses a *master* which processes the instructions and orchestrates the work. The *executors* receive the instructions from the master and perform the work.
- All instructions in PySpark are either transformations or actions. Spark being lazy, only actions will trigger the computation of a chain of instructions.

2 Your first data program in PySpark

This chapter covers:

- Launching and using the `pyspark` shell for interactive development
- Reading and ingesting data into a data frame
- Exploring data using the `DataFrame` structure
- Selecting columns using the `select()` method
- Filtering columns using the `where()` method
- Applying simple functions to your columns to modify the data they contain
- Reshaping singly-nested data into distinct records using `explode()`

Data-driven applications, no matter how complex, all boils down to what I like to call three *meta-steps*, which are easy to distinguish in a program.

1. We start by *ingesting* or reading the data we wish to work with.
2. We *transform* the data, either via a few simple instructions or a very complex machine learning model
3. We then *export* the resulting data, either into a file to be fed into an app or by summarizing our findings into a visualization.

The next two chapters will introduce a basic workflow with PySpark via the creation of a simple ETL (*Extract, Transform and Load*, which is a more business-speak way of saying *Ingest, Transform and Export*). We will spend most of our time at the `pyspark` shell, interactively building our program one step at a time. Just like normal Python development, using the shell or REPL (I'll use the terms interchangeably) provides rapid feedback and quick progression. Once we are comfortable with the results, we will wrap our program so we can submit it in batch mode.

Data manipulation is the most basic and important aspect of any data-driven program and PySpark puts a lot of focus on this. It serves as the foundation of any reporting, any machine

Another random document with no related content on Scribd:

condition:

Pork snouts,
Pork hearts,
Pork cheeks,
Pork skins,
Pork heads,
Pork hocks,
Pork ears,
Pork tails,
Beef hearts,
Beef cheeks,
Ox lips,
Sheep hearts.

These products should be thoroughly chilled by spreading them out on racks and placing them in a chill room having a temperature of from 34° to 36° F. They should be turned while being chilled. After being thoroughly chilled for from twenty-four to thirty-six hours, they should be put into vats or tierces with an eighty-degree plain pickle, using eight ounces of saltpetre to the one-hundred pounds of meat.

A wooden frame and weight is placed on the product in order to keep it immersed in the pickle. To cure these meats in vats use the following quantities of pickle:

1,400	pounds of meat will require	54	gallons of pickle.
1,000	pounds of meat will require	42	gallons of pickle.
800	pounds of meat will require	36	gallons of pickle.

The meats should be kept in a cellar during the pickling process, with the temperature ranging from 38° to 40° F., and overhauled every five, ten and fifteen days in order that all the pickle may thoroughly penetrate the meats. The different kinds of meats will be found to be sufficiently cured after being in pickle the following number of days:

Pork snouts	25 to 30 days
Pork hearts	25 to 30 days

Pork cheeks	25 to 30 days
Pork skins	10 to 15 days
Pork heads	35 days
Pork ears	10 days
Pork hocks	25 days
Pork tails	10 days
Beef hearts	25 to 30 days
Beef cheeks	25 to 30 days
Ox lips	20 days
Sheep hearts	25 to 30 days

Dry Cured Meats.—For some classes of sausage dry-cured meats are used. This consists of a process of curing meat in tierces, the meat packed closely and curing product interspersed. For this product a formula made from the following serves. For one tierce of 400 pounds use the following mixture:

16	pounds salt,
4	pounds sugar,
1½	pounds saltpetre,
2	quarts old ham pickle, which must be sweet and in good condition.

Pork and beef trimmings should be fresh, and if they have been packed in barrels for transport the blood should be allowed to drain off before being packed in the preservative. They should not be washed in pickle before being used, but should be handled dry. The two quarts of old ham pickle mentioned in the above formula should be sprinkled through as uniformly as possible when pounding the trimmings down into the tierce.

If packing fresh beef and pork hearts, head meat, beef and pork cheek meat, giblet and weasand meat, they should be thoroughly washed in a mild pickle so as to remove the blood and slime before packing in the tierce. Head, cheek, and giblet meat should not be put into ice water when cut off on the killing floor, but should be promptly removed to a cooler where the temperature is 33° to 36° F., and spread or hung up on racks to refrigerate.

Care must be taken not to allow these meats to accumulate in any bulk while warm. Hearts and large pieces should be split to reduce their size and make accessible to the curing ingredient. In the packing of these meats the pickle used with dry trimming is omitted.

Packing.—After the trimmings are properly prepared they are to be mixed with the curing ingredients. This is best accomplished by the use of a tumbler churn, weighing a given amount of the trimmings and placing with the allotted proportion of curing materials into the churn.

When mixed with the preservative, the trimmings should be put in a tierce, in layers, and pounded down as tightly as possible with a maul, and the operation continued until the tierce is as full as possible, allowing for the head to be put on. Before heading up spread a cheese cloth or thin cotton cloth over the top to protect the trimmings from the head. The tierce when headed up is removed to cold storage, where the temperature must be kept as near 38° F. as possible from thirty to forty-five days, when the trimmings are ready for use. If it is desired to keep the product sixty days, after it has been in the temperature above mentioned for thirty to forty-five days, remove to a lower temperature, 32° to 34° F.; and for more than sixty days to a temperature of 20° F.

Casings and Spices.—All classes of beef casings, namely, rounds, bungs, middles and weasands, as well as hog bungs, hog casings and sheep casings are used in the Sausage Department. There is perhaps more chicanery used in Sausage Room supplies than in any other one department, consequently care in purchase of these supplies is worthy of attention. In sheep casings it is a matter of grading as to width, pieces and yardage per bundle; in hog casings, a matter of salt per pound purchased, and grading as to width and pieces; in rounds and middles one of holes, pieces and measurement per set.

Spices should so far as possible be bought in the natural state and mixed on the premises. Pure Food laws pretty well take care of the purity of the spice in most states.

Sausage Cereals.

—This is a very important factor in the manufacture of sausage. The province of "fillers" is to absorb water, preventing shrinkage, and while this is advisable to an extent, if overdone, it detracts from the quality of the product. The main base ingredients for fillers are rice flour, corn flour and potato flour. There are many sausage fillers on the market but the foregoing ingredients are most frequently used.

Potato flour or starch is not used to any extent today, manufacturers finding that there is a great deal of trouble attached to the manufacture of sausage containing these ingredients, on account of the liability to sour and spoil. Corn flour is the best filler that can be used, being less liable to ferment, while it absorbs the water quickly. While fillers are used to a great extent, the sausage manufacturer should remember that the quality of sausage is deteriorated proportionately to the amount of water that is worked in. Hence fillers should be used with discretion, and manufacturers who aim to make a name for their goods, are frugal with fillers.

Sausage Formulas.—The following methods are tried and used for the manufacture of various kinds of sausage. Sausage makers vary procedure according to stocks on hand. However, for uniformity, it is best to conform to a standard so far as possible.

Pork Sausage.—This is produced in various grades, from a fancy breakfast quality to a substance whose chief claim to the name is the form. A good pork sausage can be made as follows:

100	lb. pork trimmings, preferably shoulder trimmings, about one-third fat.
3	pounds salt.
8	ounces pepper.
3	ounces sage.
1/4	ounce ginger.
1/2	ounce mace.

This should be chopped by passing through a 5/32 "Enterprise" plate. Mix in an arm type mixer, rather than the blade type. Mix as little as possible but sufficiently to get spice evenly distributed; stuff in medium sheep casing, 5 inch links. The matter of spicing is one of

taste and can be varied. Note the absence of water and filler in the formula.

Some makers prefer to "rock" their fancy breakfast sausage. This produces good results but is unnecessary. It is possible to use many meats in the making of this sausage and still have it passably good, but generally speaking, there is less chance for manipulation of this sort in this kind of sausage than in many of the others. The following formulas make a cheap and palatable pork sausage:

FORMULA A.

75 pounds pork trimmings,
25 pounds tripe,
8 pounds water,
3 pounds salt,
4 ounces sage,
10 ounces white pepper,
3 ounces saltpetre,
10 pounds corn flour.

FORMULA B.

90 pounds regular pork trimmings,
10 pounds tripe,
6 pounds corn flour,
10 pounds water,
2 pounds, 8 ounces salt,
4 ounces sage,
10 ounces white pepper,
3 ounces sugar,
1 pound, 8 ounces color water.

The preceding formulas are for sausage meat, often sold loose or without stuffing, also for sausage stuffed in hog casings. Stuff in medium or narrow hog casings. "Tripe" is the source of refuge to produce cheap pork sausage.

Bologna Style.—This is one of the most common and generally used type of sausage manufactured. It is in demand in nearly every locality. In the manufacture of Bologna, ingredients are used which

are not in themselves palatable, but are nutritious. The seasoning makes it palatable.

The formulas which are given below, if they are accurately followed and fresh and wholesome material carefully prepared is used, will make a sausage which is very acceptable to the consumer. This is the product that is usually made from the tougher meats such as cheeks and hearts. For a good bologna use:

25	pounds beef trimmings,
50	pounds pork cheeks,
7	pounds corn flour,
1 1/8	pounds pepper,
4	ounces coriander,
70	pounds pork trimmings,
30	pounds beef cheeks,
5	pounds salt,
2	ounces allspice,
4	ounces saltpetre,
25	pounds water.

A cheaper product can be made as follows:

45	pounds hearts (pork),
20	pounds sheep cheek meat,
65	pounds beef cheeks,
7	pounds corn flour,
4	ounces coriander (ground),
4	ounces saltpetre (ground),
20	pounds pork cellar fat trimmings,
25	pounds weasand meat,
5	pounds salt,
18	ounces pepper (ground),
2	ounces allspice (ground),
25	pounds water.

To manufacture, the product should be passed through an "Enterprise" type of grinder, using 7/64th plate. Transfer to silent cutter and chop for full five minutes, adding spice and water as the

mixture is cut and turned. Transfer to shelving room and allow to lay for twenty-four hours. Stuff in casings as required, put in smoke house at a temperature of 120° F. for one and one-half hours, raise temperature to 135° F. and carry for another one and one-half hours. Cook in water at 155° F. for thirty minutes; rinse with hot water after removal, then chill with cold water and hang in shipping room.

The length of cooking and smoking varies with the weight and thickness of the package. See schedule. This recipe is for wide middle casings.

Frankfurt Style.—This popular sausage is made from a variety of formulas. Perhaps there is no one piece of sausage as susceptible of being made so excellent or so tasteless, it being entirely a matter of what it is made from. The better grades are made from freshly killed bull beef, hashed warm. The process consists in boning bull beef and opening the meats along the seams, so to speak, skinning each bundle of muscle to remove the wrapping and cutting out all ligaments. Fresh pork, preferably, shoulder meat is treated in the same manner. The meats are then passed through a 7/64 "Enterprise" plate, and passed to a silent cutter. Here cracked ice is added in quantity and the meats cut until they are a light fluffy pulp. The spices are added during the last five minutes of cutting, and the whole mass transferred to a shelving room for twenty-four hours, when it is ready to stuff, smoke, cook, cool and sell. Wide sheep casings are used for stuffing. Make the links uniform. The proportions of meat used should be as follows:

60 pounds bull beef,
40 pounds pork shoulder (fat),
35 pounds ice,
10 pounds corn flour,
4 pounds salt,
12 ounces pepper,
3 ounces saltpetre,
3 ounces mace,
6 ounces sugar.

The following formulas are for less costly products and provide a means for disposing of some by-products:

FORMULA NO. 1.

57 pounds regular pork trimmings,
65 pounds beef cheek meat,
15 pounds cooked tripe,
25 pounds pork kidneys,
20 pounds dry salt or pickled pork trimmings,
9 pounds corn flour,
45 pounds water,
1 pound, 4 ounces white pepper,
3 pounds salt,
2 pounds color water,
4 ounces saltpetre,
3 ounces allspice,
3 ounces mace,
3 ounces coriander,
1½ ounces cloves.

FORMULA NO. 2.

90 pounds lean pork cheek meat,
60 pounds regular pork trimmings,
9 pounds corn flour,
60 pounds water,
5 pounds salt,
2 pounds, 7 ounces color water,
12 ounces sugar,
3 ounces saltpetre,
1 pound black pepper,
2 ounces mace.

The use of a mixer is unnecessary with this sausage since the silent cutter will perform this work. It should take about ten minutes to do the cutting. Smoke at 110° F. for one and one-half hours, then at 135° F. for one hour. Cook at 155° F. for eight to ten minutes, rinse and cool.

In medium and low priced frankfurts, cattle lights are used in moderate proportion, say, 10 per cent. Tripe can also be used in increased quantity.

Leona Style Sausage.—The following formula will be found acceptable for this variety of sausage:

30	pounds pork knuckle meat,
65	pounds lean pork trimmings,
50	pounds back fat trimmings or moderately fat trimmings,
22	pounds pork neck fat,
8½	pounds corn flour,
55	pounds water,
5	pounds salt,
1	pound white pepper,
3	ounces mace,
2½	ounces saltpetre,
12	ounces sugar,
2	ounces grated onions,
½	ounce garlic.

Knuckle meat to be ground through a moderately fine plate. Balance of pork should be chopped in the "silent cutter." Corn flour and seasoning should be added to the knuckle meat after it is put into the Buffalo chopper and the machine has made two or three revolutions. Chop for four minutes, stuff in eighteen-inch pieces, beef middle casings or beef bung casings. Smoke, cook as per schedule. Cool and send to hanging room.

Knoblauch Style Sausage.—Following are two formulas for Knoblauch sausage:

30	pounds pork knuckle meat,
65	pounds very lean pork trimmings,
50	pounds back fat trimmings or moderately fat trimmings,
22	pounds pork neck fat,
8½	pounds corn flour,
55	pounds water,
5	pounds salt,

1 pound white pepper,
3 ounces mace,
2 1/2 ounces saltpetre,
12 ounces sugar,
2 ounces grated onions,
3 ounces garlic,
8 ounces color water.

Stuff in beef rounds and tie with twine every five inches. Knuckle meat may be ground through a moderately fine plate. Balance of pork should be chopped in a "Buffalo Silent Cutter." Corn flour and seasoning should be added to the knuckle meat after it is put into the Buffalo chopper and the machine has made two or three revolutions.

SECOND METHOD.

50 pounds pork cheeks,
10 pounds tripe,
40 pounds standard pork trimmings,
5 pounds salt,
3 ounces mace,
9 ounces sugar,
3 ounces garlic,
35 pounds pork trimmings (lean),
15 pounds D. T. or S. P. trimmings,
9 pounds corn flour,
1/8 ounce pepper,
4 ounces coriander,
3 ounces saltpetre,
30 pounds water.

Grind pork cheeks through 7/64 plate "Enterprise" cutter. Transfer to silent cutter, add water and chop one minute, add balance of trimmings and chop five minutes, then pass to shelving room, stuff in beef round casings, tie with No. 12 twine in four-inch links. Smoke at 110° for one hour and increase temperature to 135° for one and one-half hours. Cook twenty minutes at 155° and chill after cooking, draw and pass to hanging room.

Polish Style Sausage.—Formula for making this sausage is as follows:

100 pounds beef cheek meat, or shank meat,
50 pounds dry salt or pickled pork trimmings,
50 pounds pork trimmings,
9 pounds corn flour,
30 pounds water,
1 pound white pepper,
5 pounds salt,
6 ounces saltpetre,
6 ounces coriander,
3 ounces garlic.

Grind the beef cheek meat through a 7/64th plate, add corn flour and seasoning, work in as much water as possible and then add the pork trimmings. This is a very coarse chopped sausage and the pork trimmings should be chopped about as fine as small dice. Beef is the binder of this sausage, and must be handled according to instructions. The meat, after it is chopped, can be handled the same as Bologna and Frankfurt meat by putting in a cooler for a few hours before stuffing. After the sausage is stuffed, it can also be handled as Bologna and Frankfurts, if desired, before smoking.

This sausage should be smoked carefully and strictly in accordance with the smoking schedule, as it is not cooked, this being done practically in the smoke house, during the process of smoking. After it is smoked it has a very wrinkled appearance, which is essential for this article. In fact, it is not Polish sausage unless it has this appearance.

The dicing of the meat other than the beef can be done with a rocker. Note that a "silent cutter" is not used in this manufacture. The sausage is stuffed in beef round casings.

Blood Sausage.—This sausage is made as follows:

205 pounds shoulder fat,
54 pounds pig skins,
47 pounds beef blood,

5 pounds onions,
7 pounds salt,
1 pound white pepper,
3 pounds corn flour,
8 ounces marjoram,
4 ounces cloves.

Use pickled shoulder fat and skins, cook for one hour at a temperature of 210° F. and run through fat cutting machine or cut into size of small dice. Pass the beef blood through a fine sieve in order to separate foreign matter. Cook pig skins for about two hours at a temperature of 210° F. and grind through a 7/64th plate. Mix the shoulder fat, skins, blood and seasoning thoroughly together and stuff in cap end bungs. Smoking and cooking as indicated in schedules.

Tongue Sausage.—For Tongue Sausage the following formula is given:

50 pounds hog or sheep tongue,
130 pounds shoulder fat,
34 pounds hog skins,
30 pounds blood,
8 pounds salt,
1 pound, 4 ounces white pepper,
2 pounds onions,
10 ounces marjoram,
4 ounces cloves.

Use pickled shoulder fat, skin and cook for one hour at a temperature of 210° F. and run through fat cutting machine or cut into size of small dice. Use beef blood, passed through a fine sieve in order to separate any foreign material. Cook hog skins for about two hours at a temperature of 210° F. and grind through a 7/64th inch plate. Pickled sheep tongues are preferable to pickled hog tongues, as they are smaller and make a better appearing sausage when cut. The tongue should be cooked one and three-quarter hours at a temperature of 210° F.

Before mixing the above ingredients, rinse the fat off the tongues with hot water in order to remove as much grease as possible. Mix the ingredients thoroughly with the seasoning by hand. When stuffing put about four pieces of tongue to each bung. However, this varies according to the size of the bungs used. Cap end bungs should be used in all cases. Smoking and cooking to be done as indicated in schedule.

Minced Ham.—The following formula for Minced Ham is given:

50	pounds beef trimmings,
20	pounds pork cheek meat,
80	pounds regular pork trimmings,
5	pounds corn flour,
20	pounds water,
5	pounds salt,
8	ounces sugar,
3½	ounces white pepper,
3½	ounces saltpetre.

Grind meats through a 7/64th plate; pass to "Silent Cutter," add water and spices; chop three minutes and shelve for curing. Stuff in calf bladders if available, otherwise small beef bladders.

In tying the bladders, it is best to use a wooden skewer and twine and it is preferable to use small calf bladders in place of large ones, as the time required for smoking and cooling is so long that if large bladders are used the weight of them would break the bladders where they are skewered or tied and would result in shrinkage or loss.

New England or Pressed Ham.—This ham is made from dry cured pork trimmings put down under the formula given. The best and leanest trimmings obtainable are cured for this purpose. Shoulder blade trimmings or lean shoulder trimmings are more desirable than any other kind.

After the trimmings have been cured sufficiently, which is when they show a bright cured color throughout and are without dark spots in the center of the meat, the trimmings are weighed up in

100-pound batches, and about ten per cent of lean beef trimmings, ground through a 7/64th-inch plate, is mixed thoroughly. Immediately after the trimmings are mixed the mass should be stuffed into large beef bung ends, usually from fourteen to sixteen inches long. To obtain the best results a stuffer arranged with a large sized filler is necessary. However, a hand stuffer arranged with a large sized filler, about three inches at the small end, or opening, can be used. Care should be taken to stuff the bungs as tightly as possible. They should be skewered instead of tied at the ends and should be wrapped with heavy twine, each piece having from four to six wrappings of the twine, which should terminate with a hanger for the ham. The pieces are very heavy and will break during the processes of smoking and cooking unless they are properly wrapped or tied.

This ham is smoked five hours at a temperature of from 130° to 140° F. and the house should be moderately warm before the ham is hung in the smoke. A small fire should be started to dry off the casings, after which the meat should be smoked the same as Bologna. Cook at least three hours at a temperature of 160° F. After it has been cooked it is taken immediately to a cooler, where the temperature is from 38° to 40° F., and put under a press. If no press is obtainable place the ham in layers, putting a board between each layer with a weight on top. Place the hams in a pile or under the press so that they can be picked with a long, thin skewer about one-half the thickness of a ham tryer in order to permit the water which is in the ham to escape. After pressure for twelve hours, take them out and hang up so that boiling hot water can be thrown on them to wash off the grease; thoroughly washed in this manner remove to a dry cooler.

New Jersey Ham.—New Jersey ham is made according to the following formula:

60	pounds lean ham trimmings
80	pounds lean back trimmings,
10	pounds lean beef chucks or shank meat,
4	pounds, salt,

3½ pounds cracker meal,
4 ounces formula saltpetre,
12 ounces sugar,
¾ ounce red pepper.

Beef is ground through an Enterprise 7/64th-inch plate and rocked about five minutes, when the pork trimmings are added with the seasoning. The seasoning should all be mixed thoroughly and added to the meat. The whole is then chopped about as coarse as summer sausage, or about twenty to twenty-five minutes. It is taken to a cooler after being rocked and spread about six or eight inches thick on a table, where it is allowed to remain about three days at a temperature of from 38° to 40° F.

It is then stuffed by hand stuffers into bags, which will weigh after being stuffed and dried about five pounds. These bags are made of heavy drilled cloth and should be stuffed as tightly as possible. They should be kept very clean during the process of stuffing, as any sausage meat which may stick to the cloth will leave a bad appearance after the sausage has been smoked.

After the ham has been stuffed, it should be taken to the dry room, where the temperature can be kept at all times between 46° and 55° F., 50° being preferable. The room must be airy and dry and it will take at least ten days under favorable circumstances to get the ham in proper condition to smoke. It should be smoked about four hours in as cold a smoke as possible, 70° to 75° F. being as hot as it is safe to smoke it, 60° F. being nearer the proper temperature. After it has been smoked, it should be again hung in a cool temperature for three days, when it will be ready for shipment. This sausage is manufactured extensively in New Jersey and the east.

Head Cheese.—Head cheese is made as follows:

44 pounds cooked pig skins,
55 pounds cooked pig snouts,
33 pounds cooked pig ears,
55 pounds cooked beef hearts,
51 pounds cooked neck fat,

20 pounds water in which the meat has been cooked,
1 pound white pepper,
10 pounds onions,
4 ounces allspice,
2 ounces cloves,
3 ounces marjoram,
3 ounces carroway seeds.

The cooked meats are chopped by hand with a knife until reduced to the proper size, except the skins, which are ground through a 7/64th-inch plate after being cooked. The mass usually is mixed by hand and stuffed into cured hog paunches or beef bungs and cooked as per cooking schedule appended hereto. After the sausage is cooked, it is taken to a cooler and usually pressed by laying the paunches or bungs side by side with a board between each layer and a moderate weight on top of the last board. However, if properly made this is unnecessary as the gelatine from the skins and the water in which the meat has been cooked will bind the other ingredients together sufficiently without much, if any, pressing.

Boneless Pigs Feet.—This product is prepared as follows:

25 pounds fresh pigs feet,
30 pounds fresh pigs skins,
15 pounds fresh pigs snouts,
15 pounds fresh pigs ears,
20 pounds fresh pork trimmings,
15 pounds fresh beef trimmings,
10 pounds white pepper,
50 pounds water in which meat has been cooked,
4 pounds, 1 ounce salt,
4 ounces cloves.

Use one gallon (45-grain) vinegar to five-hundred pounds of the above mass. Cook all of the meats in one vat, thoroughly, in pudding nets, and chop same as headcheese, mix seasoning, water and vinegar with the meat in a large tub or tight-bottom truck.

It is necessary to use tin moulds for this sausage and they are generally of one size, shaped as a ten-pound wooden bucket or other sized packages which may be intended to be used for shipping purposes. Fill these molds with the mixed mass and put on top of each a wooden block the size of the mold and about three inches thick. Then remove to a cooler and press tightly by placing on top a board with a weight. In order to obtain the best results, the molds or cans should be cooled quickly, therefore a temperature of about 36° F. is desirable. To remove the contents from the cans or molds, submerge in hot water for a few seconds, when the meat will loosen from the sides of the molds and can be turned out readily.

After the product has been removed from the molds allow it to stand for a short while in the cooler before placing in shipping packages.

This sausage can be made without using wooden tops on the cans or molds and without pressing it. If the pig skins, after they are cooked, are ground through a 7/64th-inch plate and then mixed with the mass, more of a jelly will be produced and they will not require pressing. In preparing meats be particular to remove all bone, gristle or cartilage.

Liver Sausage.—The following formula is for Liver Sausage:

20	pounds cooked lean pork trimmings,
20	pounds cooked pork cheek meat,
20	pounds cooked pork skins,
10	pounds cooked hog livers,
50	pounds cooked tripe,
6	pounds cooked shoulder fat,
3	pounds salt,
3	pounds onions,
9	ounces white pepper,
2	ounces marjoram,
2	ounces cloves,
1½	ounces allspice.

Above is all ground through a 7/64th-inch plate except the shoulder fat, which is run through a fat cutting machine or cut into

size of small dice. It is necessary to mix this sausage in a sausage mixer. The seasoning should be put into the mixer when starting to mix, but the shoulder fat should not be put in until about half through. Stuff immediately into hog bungs, or beef middles, as desired. Cook immediately as per cooking table and then place in cooler, at a temperature of 36° to 40° F. until thoroughly chilled, when it is ready for shipping.

Boneless Ham.—This is made from pork shoulder butts, cured in sweet pickle and stuffed in small No. 2 beef bungs. Smoke forty-eight hours at a temperature of 120° F. The bungs may be slightly colored, the same as Polish sausage casings, before stuffing, if desired. Not cooked.

Cottage Ham.—This is made from boneless ham butts or shoulder butts, cured the same as boneless ham butts. It is not stuffed but strung from the large end of the butt and smoked thirty-two hours at a temperature of 120° F. and not cooked.

Stuffed Hogs Heads.—Select a well shaped head, cut off about three to four pounds behind the ears and remove the bones. Care should be taken in entering above the eyes, where the skin is thinnest and lies directly on the bone. Do not remove the snout bones, but saw off the hindmost jaw bone right behind the mouth. Remove the cheek meat on either side until with the skin it is about one-half inch thick. Cut off about three inches square from the lower cheek at the back to make the head more shapely. Sew from the snout up the back bone, where the head is to be filled and cut around cover from the skin to fit the back opening. Prepare the stuffing as follows:

Good firm young pork, moderately fat, is coarsely chopped with the required quantity of salt and allowed to stand twenty-four hours. Use about ten pounds chopped fine, spiced with five and one-quarter ounces of salt, one-half ounce of ground white pepper, seventy-seven grains of finely ground mace, mixed with one pound of boiled tongue meat cut into shape of dice, one-half the size of a walnut. Mix all thoroughly and fill head with the mass. Sew the cover on and smoke for two hours until it turns to a yellow brown color.

After it is smoked, tie the head up in a cloth, wrapping a string around it very evenly from front to rear. Cook in boiling water from three to three and one-half hours, and then allow to cool. The head must be a chestnut brown when thoroughly smoked and cooked.

Scrapple.—For making scrapple use two pig heads, two pig tongues and two pig livers. These should be cooked in an iron-jacketed kettle that will hold about forty-five gallons. Cover thoroughly and then remove from kettle and cut up the same as for head cheese. After the heads, tongues and livers are taken from the water, skim the grease and add forty pounds of corn meal and five pounds of buckwheat to the water, putting in a little at a time, stirring as put in. Cook slowly for five hours. Seasoning should be added before buckwheat and corn meal are put in, consisting of:

2 ounces white pepper,
1 ounce red pepper,
8 ounces sage,
4 pounds salt.

After the meal and water has been cooked about four hours and forty-five minutes, add the heads, tongues and livers, stir thoroughly about fifteen minutes, shut off steam, and place in pan. When about cool, but while plastic, add to each pan top a coating of the grease skimmed from cooking.

Bologna Sausage in Oil.—This is a sausage manufactured quite extensively by packers, who find the principal markets for it in the south. In fact, it is not used anywhere but in warm climates and it is usually put up in twenty-pound and fifty-pound tin packages.

Much experimenting has been done to ascertain the best size for packages to put up sausage in oil. The following formula is generally considered to be a good method for making this sausage:

20 pounds fresh head pork meat,
50 pounds fresh pork hearts,
30 pounds fresh regular pork trimmings,
15 pounds fat pork trimmings,

80 pounds fresh beef cheek meat,
1 pound, 8 ounces corn flour,
3 pounds, 8 ounces salt,
1/2 ounce cloves,
1/2 ounce coriander,
3 ounces saltpetre.

Stuff in different sized beef rounds. The beef cheek meat, pork hearts, and pork cheek meat are ground through a 7/64th inch plate, and afterward chopped with a "Silent Cutter" the seasoning being added at the same time.

Use no water in this sausage under any circumstances. After the beef and beef hearts, also the pork cheek meat, have been chopped as fine as desired, add the pork trimmings and chop the same as any other Bologna.

It is desirable to stuff this sausage immediately after it is chopped, and if a steam stuffer is used care should be taken that no water from the evaporation of the steam is allowed to get into the sausage. The bench where the sausage is stuffed should be absolutely free from water or moisture. This is the principal factor in the successful manufacture of this product.

After the sausage is stuffed, it is smoked about three hours at a temperature of 150° to 160° F., or until it is dry clear through. This sausage is not cooked. Keep it away from all water and moisture.

After the sausage is smoked allow it to cool in a dry airy room, but do not put it in a cooler. When it is thoroughly cool, pack into twenty-pound and fifty-pound packages, as desired, as follows: In twenty-pound cans, place sixteen pounds Bologna and four pounds oil. In fifty-pound cans, place thirty-six pounds Bologna and fourteen pounds oil.

In order to pack the cans properly, it is necessary to stuff different sized beef rounds, as mentioned above, so that they will fit in nicely without breaking the casings, and without filling the cans too full.

After the cans have been filled with the required amount of Bologna, crimp on the summer top, which has a two-inch hole and a cap to fit. Fill the cans as full as possible with deodorized cotton

seed oil, which must be cold. Allow the cans to stand for thirty minutes, then refill so that the oil runs over the top through the hole, put on the cap immediately and solder right through the oil which will accumulate around the cap and on the top of the can. This will not hinder the process of soldering and it prevents the possibility of any air getting into the cans.

After the caps have been secured, solder around the crimps of the summer top. This can be done before the oil is put in if desired. Extreme care must be used in soldering the cans so that no air whatever gets in, or oil leaks out, as the sausage will spoil if this occurs.

The cans should also be fitted with the regular covers so as to protect the summer top. Pack in crates, the twenty-pound size, two to four to a crate; the fifty-pound size, one to two to a case.

The case should be large enough so as to admit of packing sawdust beneath the bottom, around the sides and on the tops. A crate large enough to permit one-half inch space around the cans is the size generally used and there should be a partition in the crates where more than one can is packed in a crate.

In freighting this class of merchandise in the south, in fact wherever it is shipped, it receives more or less rough handling and a great many freight handlers use box hooks, which they stick into the sides of the crates, and if there is not sufficient protection from the amount of sawdust put in, the cans are punctured, the oil leaks out and the sausage spoils. Sausage handled in the above manner has been known to keep two years in temperature ranging from 40° to 100° F.

Pork Sausage in Oil.—Use the same formula as for regular pork sausage except that the trimmings must be moderately lean, and the sausage *absolutely* free from water. Use also the same spices. Stuff immediately after the sausage is chopped, using the same care as to moisture as for Bologna in oil, and smoke over a very cold smoke until the sausage seems dry all the way through. Allow it to cool, handling and packing in every other respect the same as Bologna Oil.

Bologna Varnish.—Where it can be used and not conflict with Food regulations bologna varnish can be used to advantage. It improves the appearance and decreases shrinkage. As considerable water is used, in addition to the natural moisture in meats there is, after it is manufactured, considerable shrinkage, and if allowed to hang for any length of time, the casings become wrinkled on account of the evaporation of this internal moisture. The varnishing of the sausage creates a covering on the outside which prevents this wrinkled appearance, improving its looks, and retarding the evaporation of moisture from the sausage. The formula is as follows:

6 pounds white shellac,
1 pound boracic acid,
2 pounds aqua ammonia,
14 pounds of water.

The mixture should be put into a vessel and heated to a point where the shellac is well dissolved. When this is accomplished, add four gallons of water. This varnish, in order to be ready for use at any time, must be kept lukewarm. It should, therefore, be kept in a jacketed pan, surrounded by either hot water or steam, to hold it at the proper temperature. The sausage should be immersed and immediately hung up to dry.

This varnish can be used without any detrimental effect whatever on all kinds of smoked Bologna sausage or smoked cooked pressed ham. It preserves the sausage, keeps it from molding, and is especially effective where it is necessary to pack Bologna in boxes for shipment long distances.

Where the dipping pan is used, it is only necessary to dip the sausage in the above solution a few seconds before hanging on racks to dry. The sausage is usually ready for shipment in one hour after it is dipped, if the preparation is properly made, where a large amount of Bologna is being dipped it is, of course, necessary to have a larger dripping pan and a larger quantity of varnish. The proportions should be increased accordingly.

It is also important that, after the varnish has stood from one period of dipping to another, to skim the grease off the top of the varnish before again using it, and the Bologna should always be dipped immediately after it is taken from the cooking vats; in other words, while *hot*.

Boiled Ham.—Boiled ham and shoulders, also cooked meats, are usually included as a sausage product. There is nothing that determines the cost of the finished product as much as the shrinkage, hence the method that will produce the least loss in weight from original to finished product is the process desired.

There are two methods of cooking hams, one is to steam them in a retort or some receptacle where they are cooked by the heat generated by steam; another is to cook them in water. The latter process, from careful observation, seems to be the one that gives the best results as regards the shrinkage, although steaming perhaps makes the ham more palatable. The hams before being cooked should be bound and wrapped with twine to hold them in shape. A form made of galvanized iron, clamping plates which are put into a press, with the ham tightly clamped on the inside should be used. The ham is cooked and chilled in this mold.

The cure of the ham has much to do with the shrinkage and it is therefore preferable to use fully cured hams instead of old cured hams, as the shrinkage is greater on over-cured meats. It is also advisable to sort the hams as to size, having each vat or tank of hams uniform. If not uniform in size there is an excessive shrinkage on small hams which are overcooked. In all cases the hams should be soaked, removing the surplus salt. The length and time of soaking depends altogether on the age of the meats. The hams should be thoroughly washed and if they are to be branded this should be done before they are boned or cooked.

Rules for Boiling Hams.—When hams are boned (if desirable) and wrapped, they should be put into a vat of water, temperature about 212° and the temperature regulated until it reaches 155° to 160° F. The hams are held at this temperature until they are cooked, which requires somewhat longer time than when they are cooked, by

steam at a higher temperature. A twelve-pound ham will require from four and one-half to five hours. After the hams are cooked they should be allowed to cool in the water in which they were cooked; not taken out, or drained, or set in the cooler, for in the water in which they are cooked are juices which are absorbed by the hams as they cool, and the shrinkage is much less than if taken out immediately. The hams should then be taken to the smoke house, laid on racks and given a very light smoke, then to the cooler.

Shrinkage in Boiling Hams.—Hams taken out of pickle and drained for twelve hours to shipping weight, will show the following shrinkages under favorable circumstances:

	Per cent.
Hams not boned, smoked after cooking	9 to 12
Hams with bone out, including the shank bone, skin on, not fatted	12½ to 18
Hams with bone out, skin lifted, fat removed	18 to 23
Hams with bone out, the skin and fat removed	33 to 40
Hams skinned, fatted, bones left in	28 to 35
Skinned shoulders, bone out	30 to 35

CHAPTER XXVI
DRIED SAUSAGE

SUMMER SAUSAGE — PRESERVATIVES — COOLING ROOM — STUFFING — HANGING ROOM — SMOKE HOUSE — DRY ROOM TREATMENT — DRY ROOM CAUTION — SHIPPING AGES — STORAGE — PREPARATION OF CASINGS — TRIMMING MEATS — FORMULAS FOR SAUSAGE.

Summer Sausage.—Under the head of "Summer Sausage" we take up an entirely different article, a sausage that is dried and smoked (not cooked), during which time the ingredients used for seasoning effect a cure. This sausage will keep for months if properly handled. It is necessary, however, that every detail be very carefully watched, as a slight omission or error in its manufacture causes immense losses at times.

As this is strictly an air dried sausage, weather conditions have a great deal to do with its successful manufacture, and it is necessary to have special facilities in the way of coolers, smoke houses and dry rooms, the proper arrangement of which is somewhat expensive. Air conditioning and fan circulation can be used. It is impracticable for any manufacturer to attempt to make this article in large quantities unless he has proper facilities, and it is the manufacturers who do make it in large quantities that are financially successful, and are able to keep their cost of production to a minimum.

Preservatives.—Previous to the enactment of the Pure Food Laws the liberal use of preservatives made the handling of summer sausage, while always requiring care, a matter of not so great consequence as at present. Now, the only preservative used is the agent saltpetre, and such preservative help as is obtained from the spices which are used. These conditions require the use of absolutely sweet materials.

Cooling Room.

—Clean airy cooling rooms are necessary. They should be provided with spreading shelves of sufficient area to allow the cut spiced meats to be spread from twelve to forty-eight hours so as to enable the spices, saltpetre and salt to permeate the product before stuffing and to make the meat firm. If the sausage is stuffed immediately after it is rocked, it is liable to wrinkle in the smoke house, which gives it an unsavory appearance. Temperature of room should be 37° F.

Stuffing.—The stuffing should be performed in a semi-cool room, using mechanical or hand operated stuffers at a moderate pressure. Some manufacturers use steam or hydraulic equipment, but these are usually arranged so that no moisture, condensation or drip comes in contact with the meats. The temperature of the room should be about 45° F.

Hanging Room.—A hanging room in which the temperature can be properly controlled should be available in which the sausage can be hung until the casings are sufficiently dried for smoking. While the sausage should be dry, it should not be allowed to become so dry that the casings are glossy or hard. This is a very important matter as the sausage will not take the correct color if the smoke does not penetrate the casings, or if they are allowed to become too dry. This applies particularly to sausage stuffed in beef middles, also to sausage stuffed in hog bungs, though not to so great an extent.

In preparing the sausage for the smoking process, by endeavoring to prevent the sausage from becoming too dry there is danger of their becoming slimy, which is more detrimental than excessive dryness. Slimy sausage will not take the smoke and will sour quickly if not properly handled. Sausage in both beef casings and hog casings should be dried before smoking, so that the outside will feel about dry to the touch.

In order to obtain this result it is necessary to have plenty of hanging room so that the proper temperature and ventilation is available for the entire lot of sausage which is being prepared for smoke. If the sausage is too closely hung it should be moved about

from the center to the sides of the room occasionally so that all the sausage receives the same ventilation.

The manufacturer should have sufficient space to allow the sausage to hang after it is stuffed and before it is smoked, for from two to three days, and sometimes longer. In the winter season, which is the proper time for manufacturing high grade summer sausage, it is a good practice to allow it to hang as long as possible before smoking, but it must be watched to prevent sliming or becoming too dry.

Smoke Houses.—The smoke houses should be built of brick, sheet iron houses which have been experimented with, especially for summer sausage, have proved to be complete failures. The brick not only protect the house from the varying outside temperatures, but retain the heat, which is desirable and necessary in the successful smoking of this kind of sausage.

Smoke houses are preferably equipped with rails and trolleys. The houses are usually built like ham houses, viz: 12 × 12 feet, or some size convenient to the trolleys to be used.

The distance from the fire should be in any case, whether the track system or the ordinary smoke house with the beam system is used, twelve feet, and in some cases, such as in smoking summer sausage in bladders, or Braunschweiger in hog casings, the distance from the fire should be twenty feet or more. This, however, will be explained in the formulas for making the different kinds of sausage.

The main point to be considered in the construction of a smoke house for summer sausage is to have it so arranged that the heat can be regulated to different temperatures, also the amount of smoking, as each kind of sausage requires a different temperature. The ventilation of the smoke house should be perfect and absolutely controllable, as the weather conditions have a great deal to do with the successful smoking of sausage and the houses should be arranged so that they can be kept at a uniform temperature and humidity during any kind of weather.

As in the case of domestic sausage, no smoke house should be used for summer sausage unless the temperature can be maintained

uniform and the walls in proper condition.

Since, summer sausage is not cooked before it goes to smoke, a cold or damp smoke house will "ring" the sausage even quicker than a cold smoke house will "ring" Bologna, therefore extreme caution should be used in this particular.

Hardwood and hardwood sawdust are used exclusively in smoking summer sausage, and both wood and sawdust should be absolutely dry when the fire is started. In some cases, however, after the sausage is very nearly smoked, it is advisable to use a little damp sawdust before completing the operation.

Dry Room Treatment.—After the smoking process is completed the sausage is taken to the drying rooms where the temperature can be kept at all times between 46° and 53° F., the proper temperature being 48° F., if it can be maintained. The dry room must be fitted with steam pipes running underneath the sausage and around the sides of the room and underneath the windows in order to supply the necessary heat. The room must be supplied with ample windows for light and ventilation and should be very high so as to permit the required overhead ventilation. At all times the windows must be kept open a little to allow fresh air to enter no matter how cold the outside temperature. If the weather is damp the windows nearest the top or the top ventilator of the room should be opened a little. Sausage is not usually hung adjacent to the windows.

Steam should always be turned on in damp weather to dry the air, providing the weather is not too warm and the temperature in the room can be kept as low as 53° F. The room should be arranged in sections, so that there may be an empty section between each lot of new sausage. As the sausage becomes drier it can be hung more closely. The sausage, should not dry too quickly as too much air will dry it near the casing, which will cause the sausage to stick to it and become dry. In that case the inside will not dry uniformly and the sausage will wrinkle and in some instances become sour.

The different kinds of sausage require different places in the dry room. Some require an abundance of air and others, like "Holsteiner" and "farmer" sausage, if properly smoked, can be hung

where it would not be policy to hang summer sausage in hog bungs. As both of these sausages are coarse chopped, they can be handled with much less fear of being spoiled than the finer chopped sausage. However, with every description of dry sausage, constant attention must be given or poor results will follow.

Summer sausage in hog bungs can be subjected to more draft or air than summer sausage in beef casings. Consequently beef casings are generally hung near the center of the room where they receive plenty of air, but no drafts. This is a place where the human element comes greatly into play.

Dry Room Caution.—Do not hang green and dry sausage in the same room. It is advisable to keep dry rooms for smoked sausage as free from mold as possible. While a slight mold does not hurt summer sausage (in fact some summer sausage requires this before it is ready for shipment), it will be found that smoked sausage drys better and quicker in a room that can be kept free of mold. Sausage that molds too much before it is dry necessarily has to be washed. This process does not hurt the sausage, and in some cases washing does it good, especially if by neglect or otherwise it has become greasy in the smoke house. Sausage will not dry as rapidly if greasy and the process of washing it quickens the drying. In washing sausage warm water, not hot, should be used. A little sal soda in the water is desirable.

Shipping Ages.—Summer sausage in both hog bungs and beef casings if properly handled can be shipped, in three stages of dryness, as follows: New, twenty to twenty-five days old; medium dry, forty to forty-five days old; dry, sixty to seventy-five days old. In cheaper grades of summer sausage, many kinds of which are manufactured, can be shipped in much less time than indicated above. In fact, there is sausage made which can be shipped almost immediately from the smoke house. This sausage is allowed to stand for some time after chopping and before stuffing, previous to being put in the smoke house. It is then smoked very hard, or with more heat than the better grades of this sausage. Some manufacturers

use more heat than smoke, but it does not produce a first-class article.

Storage.—Summer sausage when thoroughly dried should be stored in a cooler at a temperature of 35° F. It can be so handled if boxed, but it is preferable that it hang from racks, closely, since it will require an inspection upon shipping.

Preparation of Casings.—In preparing casings for summer sausage of all kinds it is necessary, in order to insure good results, that casings be prepared, certainly hog bungs, at least thirty days or even several months before they are used. There are always many fat bungs in hog casings and in preparing them and putting them down in salt brine for thirty days or longer, the fat on the casings becomes dry and there is less danger of the sausage becoming sour. Summer sausage will become sour as quickly from using fat hog bungs as from any other cause, therefore these must be thoroughly fatted before they leave the preparing rooms. They are generally in good condition in this respect when received, and therefore do not require as much time in the curing or preparing as do hog bungs. Fat beef middles or beef rounds spoil the appearance of the goods.

Trimming Meats.—The selection and the trimming of meats for summer sausage is a matter of great consequence. Sinews, muscle cover and tough binding elements should be removed so as to avoid the sausage being tough in the eating. All knives and cutting tools should be sharp.

Formulas.—Under the formulas given below are descriptions of the methods of manufacture in detail rather than generalizing. Where the term "rocking" is used it describes cutting with a rocker cutter. The description for making Cervelat should be carefully noted, as the methods are utilized with slight changes in making other kinds.

Cervelat Style Sausage.—A very satisfactory formula for this sausage is as follows:

FORMULA A.

40 pounds beef chucks, very lean and entirely free from all sinews,