|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  | | --- | --- | --- | |  |  |  | | МИНОБРНАУКИ РОССИИ | | | | Федеральное государственное бюджетное образовательное учреждение  высшего образования  **«МИРЭА** – **Российский технологический университет»**  **РТУ МИРЭА** | | | | |
| Институт информационных технологий (ИТ) | |
| Кафедра прикладной математики | |

|  |  |
| --- | --- |
| **ОТЧЁТ ПО ПРАКТИЧЕСКОЙ РАБОТЕ № Номер практики** | |
| **по дисциплине «Большие данные»** | |
|  | |
| Выполнил студент группы КМБО-07-22 Невский В.Е. | Горячев А.А |
| Проверил ассистент кафедры ПМ ИИТ | ФИО Семинариста |

Москва 2024

**Оглавление**

[**Задание 1** 2](#_Toc160121565)

[**Задание 2** 4](#_Toc160121566)

[**Задание 3** 5](#_Toc160121567)

[**Задание 4** 7](#_Toc160121568)

[**Задание 5** 8](#_Toc160121569)

[**Задание 6** 10](#_Toc160121570)

[**Задание 7** 11](#_Toc160121571)

[**Задание 8** 13](#_Toc160121572)

[**Задание 9** 16](#_Toc160121573)

[**Задание 10** 18](#_Toc160121574)

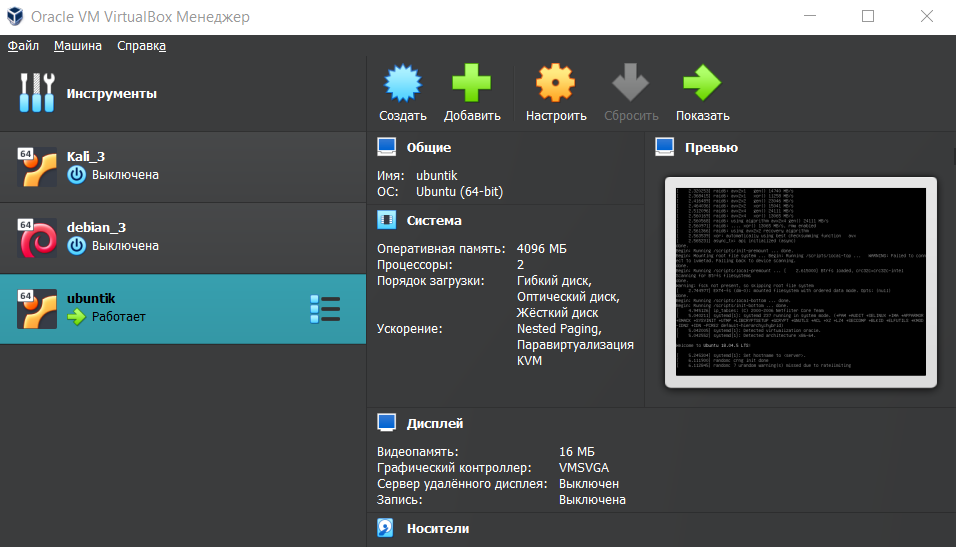
[**Задание 11** 19](#_Toc160121575)

# **Задание 1**

* Развернуть операционную систему Ubuntu Linux Server 18.04 с помощью средства для виртуализации ОС VirtualBox на персональном компьютере.

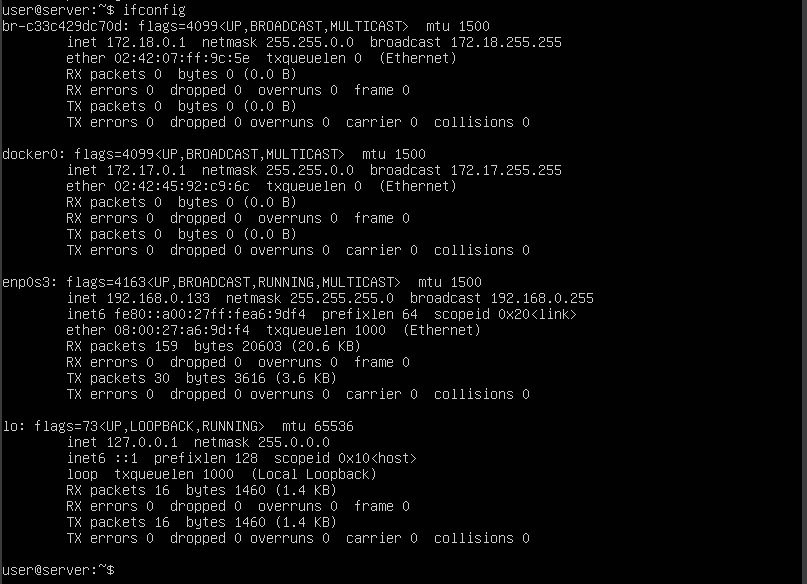
Запустим VirtualBox и импортируем конфигурацию, заранее скачанную по ссылке: <https://drive.google.com/file/d/1pXbcGvuIHF6CXwVY0406RSfx1N_3mJqn/view?usp=drive_link>

Настроив необходимое нам количество оперативной памяти и процессоров наша ОС появляется в списке:

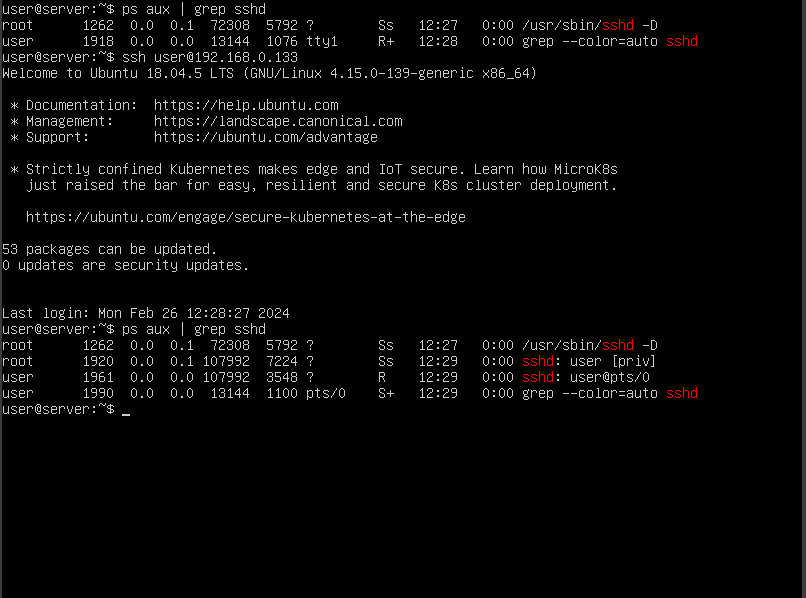


* включить серверную ОС из системы виртуализации, войти в систему, узнать ip адрес машины с помощью утилиты ifconfig; подключиться к серверной ОС с помощью терминала командной строки пользовательской с помощью утилиты ssh по полученному ip адресу сервера.

Запускаем нашу виртуальную ОС. Входим с помощью логина: **user** и пароля: **user** в систему. Далее нам необходимо узнать ip адрес машины. Воспользуемся утилитой **ifconfig**:



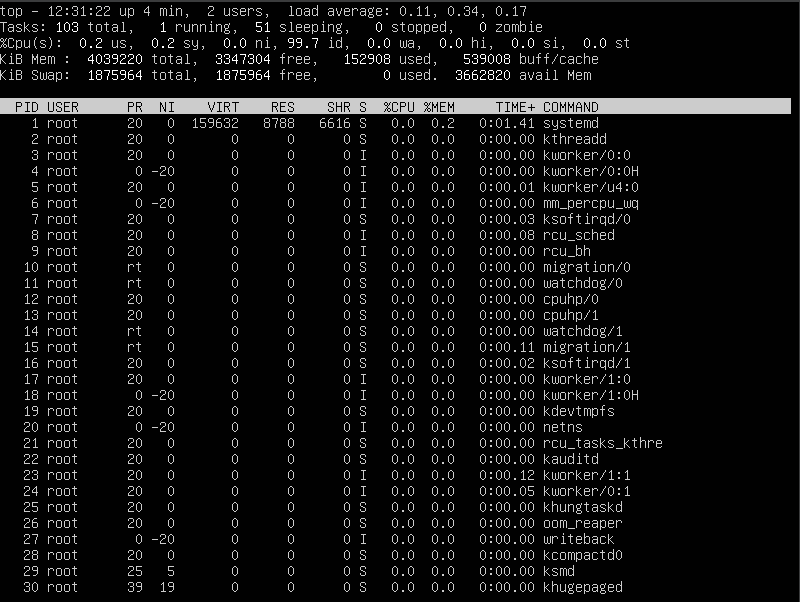
Теперь, используя **ssh** подключимся к серверу:



# **Задание 2**

Получить справку по команде top. С помощью команды top просмотреть занимаемое операционной системой место в оперативной памяти. Выйти из выполнения команды top в терминале.

Просто введём в терминале команду **top**:



Как мы видим, наша виртуальная ОС занимает около 150 000кб.

Выйдем из выполнения команды top нажатием клавиши ‘q’.

# **Задание 3**

Просмотреть путь к текущей директории. Создать в пользовательской директории папку для самостоятельной работы цифры\_шифра/my\_test\_folder. Внутри нее создать директории в связи со специальной структурой (рис. 1). Вывести на экран всю созданную древовидную структуру в виде списка папок и подпапок в терминал командной строки. Перевести вывод созданной структуры в файл ~/my\_test\_folder/temp/dirs.txt.



*Рисунок 1*. Структура директорий для создания в my\_test\_folder.

Посмотрим путь к текущей директории командой **pwd**, а затем создадим предложенную нам структуру пользуясь командами **mkdir**.



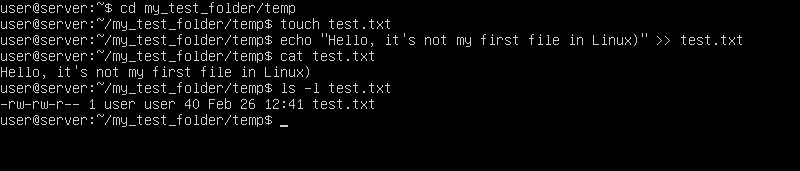
Переведём вывод данной структуры с помощью команды **tree** в файл dirs.txt. Прочитаем вывод с помощью команды **cat**.



# **Задание 4**

Перейти в папку /my\_test\_folder/temp. Создать в папке пустой файл test.txt. Ввести в файл информацию «Hello, its my first file in Linux!». Вывести сообщение из файла в консоль. Просмотреть размер созданного файла.

Переходим в нужную нам директорию с помощью **cd** и создаём test.txt пользуясь командой **touch**. Дальше просто записываем текст командой **echo.** Смотрим размер файла (в байтах) с помощью **ls –lh**.



# **Задание 5**

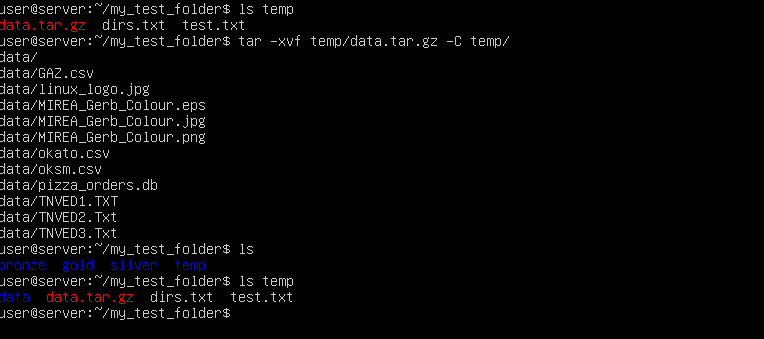
С помощью команды wget загрузить в папку temp файл архива https://github.com/qwerty29544/BigDataEssentials/raw/main/Practice1\_LinuxCom

mands/data.tar.gz с данными для выполнения дальнейших пунктов практической работы. Разархивировать файлы в папку temp, не создавая новых каталогов. Просмотрите список файлов и их уровней доступа.

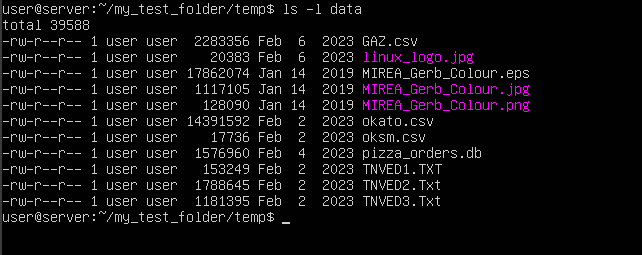
Скачиваем архив пользуясь **wget**.



После успешной установки распакуем архив командой **tar –C** (без создания поддиректорий).



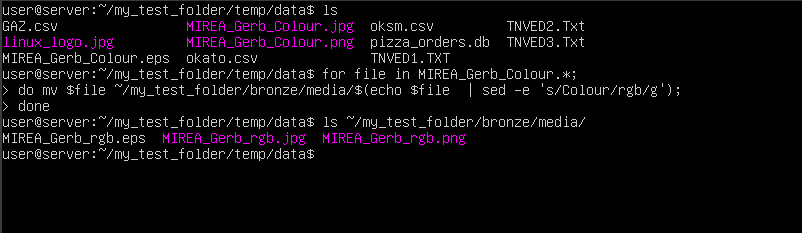
Посмотрим на наши файлы и доступ к ним.



# **Задание 6**

Перенести все файлы изображений (расширения .eps, .png, .jpg) в папку /bronze/media. Вывести список файлов в данной директории, продемонстрировать количество занимаемого места на диске данными файлами, а также список прав на доступ к файлу. Переименовать изображение герба РТУ МИРЭА в MIREA\_gerb\_rgb.\*

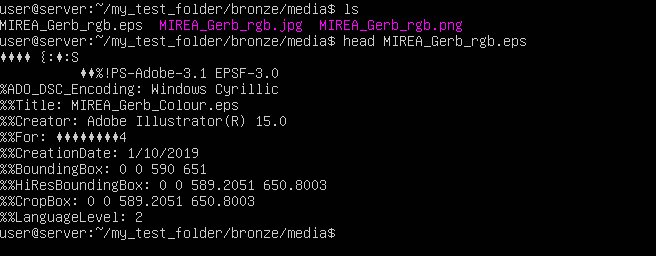
Пройдёмся в цикле **for** по нашим картинкам и каждую переместим с переименованием.

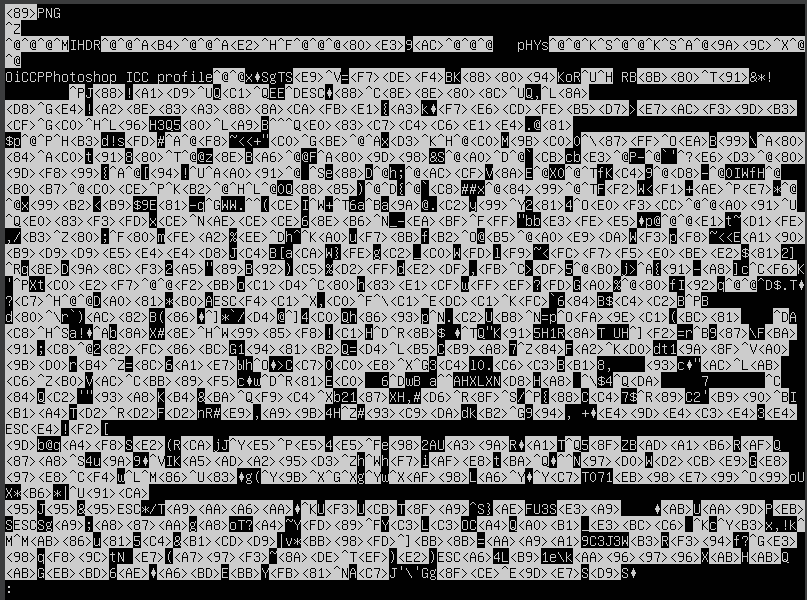


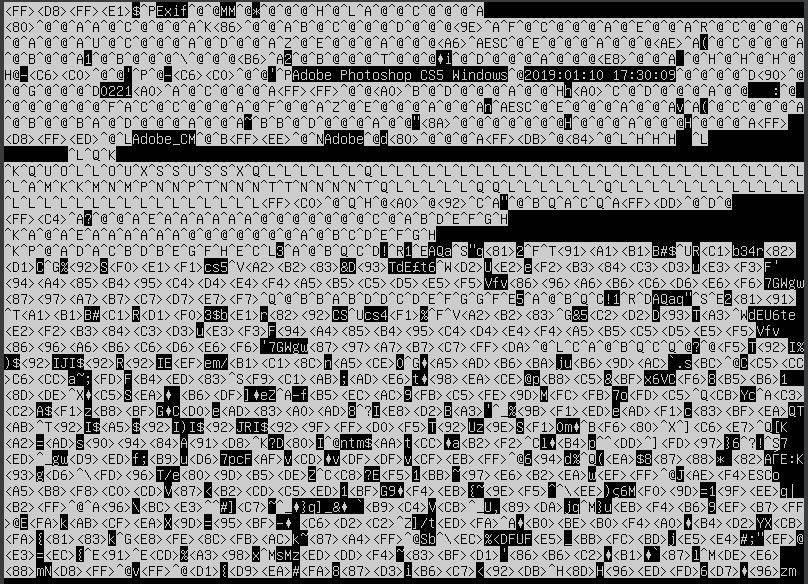
# **Задание 7**

Вывести в терминал первые строки из файлов изображений. Показать какой формат изображений выводит информацию в структурированном виде. Чем визуально различаются префиксы во всех форматах файлов изображений.

Команда **head** выводит первые строки файлов.







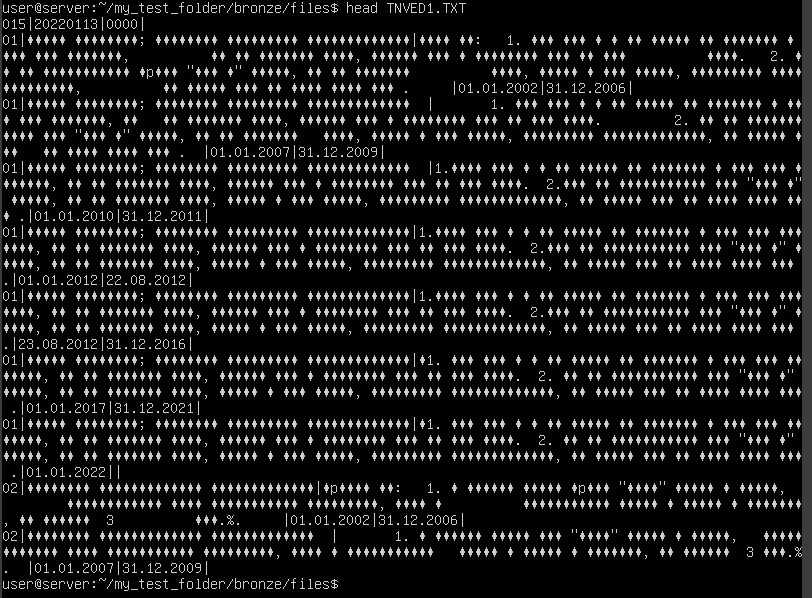
Вот в таких форматах выводятся данные изображений.

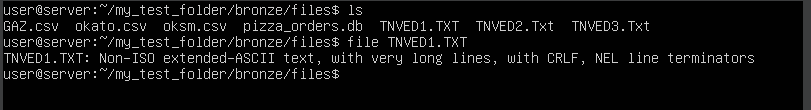
# **Задание 8**

Перенести файлы форматов .txt, .TXT, .csv, .db в папку /bronze/files. Для файла TNVED1.TXT вывести первые его строки. Проверить кодировку файла TNVED1.TXT с помощью команды file. Создать новый файл с именем tnved1\_utf.txt, переведя кодировку файла TNVED1.TXT из кодировки CP866 в кодировку UTF-8. Вывести ещё раз содержимое файла в терминал. Проделать то же самое с файлами TNVED2 и TNVED3. Перенести файлы в читаемой новой кодировке в папку /silver/files.

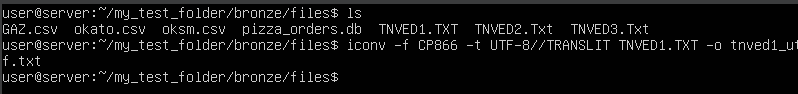
Перенесём файлы, посмотрим первые их строчки и проверим кодировку.

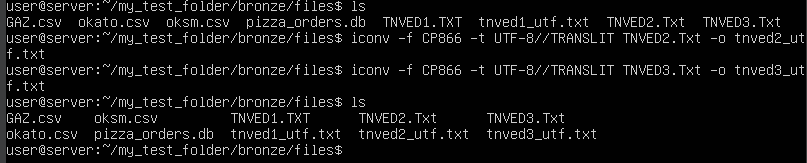




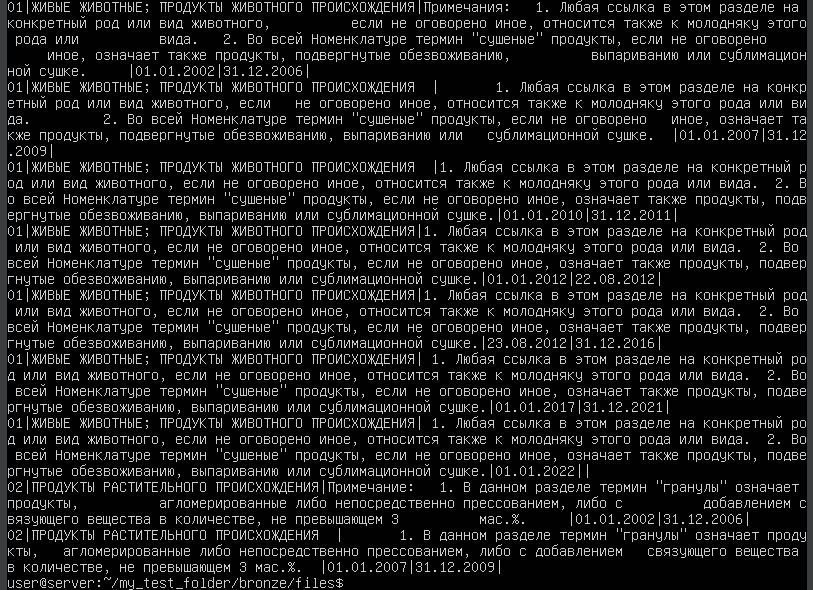


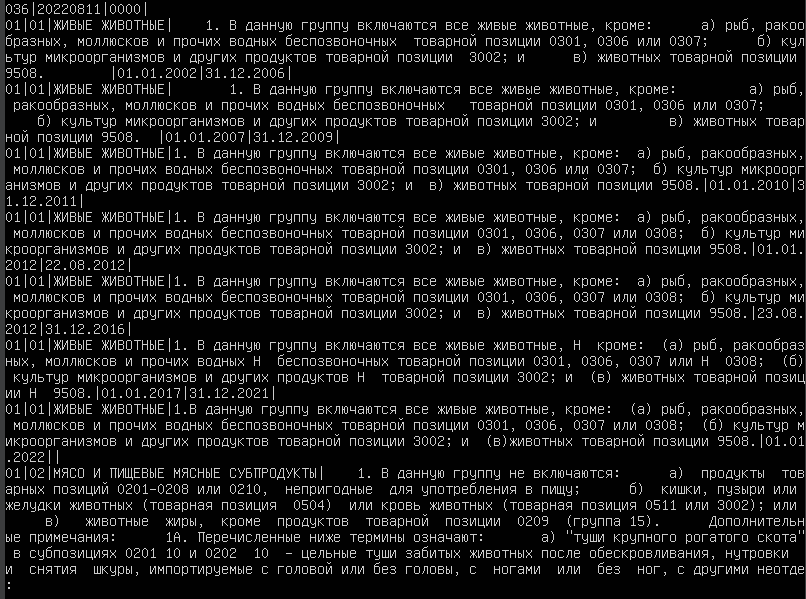
Переведём файлы в кодировку UTF-8.

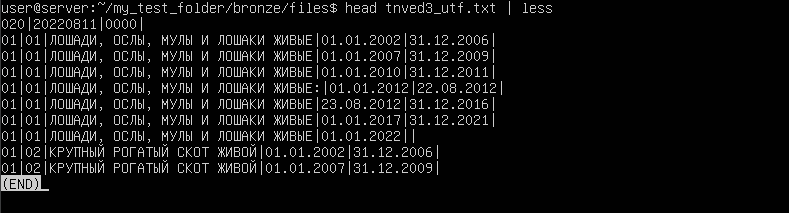




И посмотрим на итог.

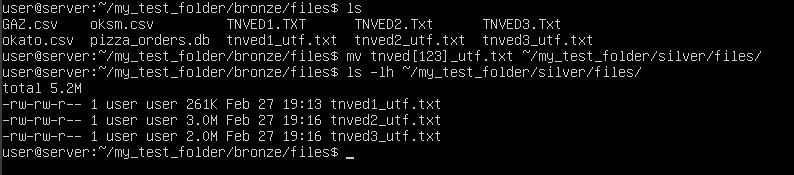






Как видим текст стал читаемым для нас.

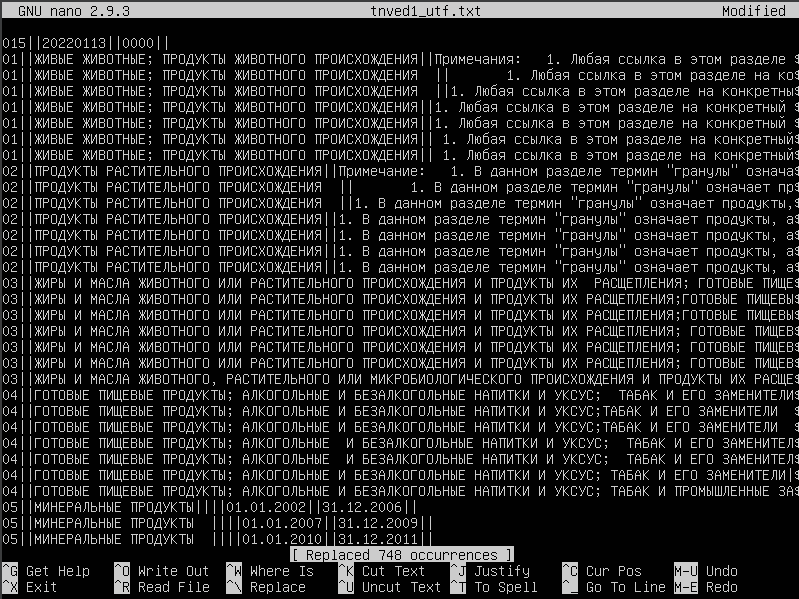
Осталось только перенести файлы.



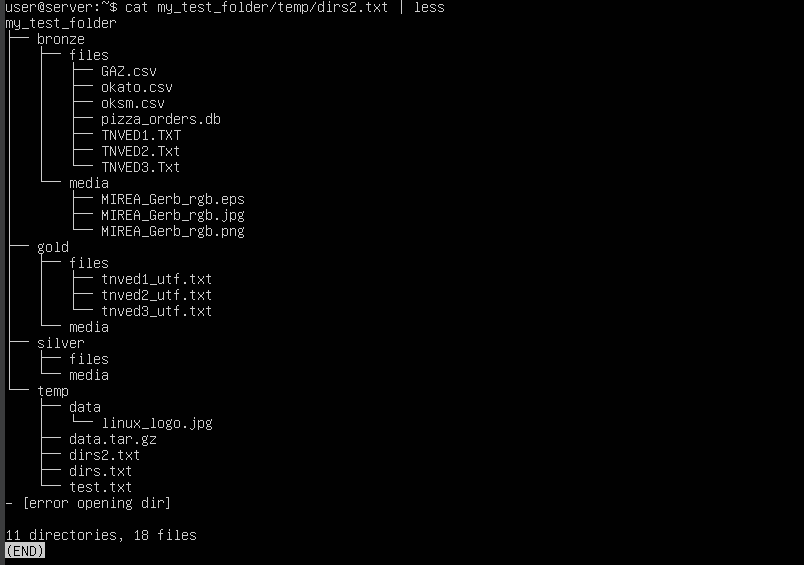
# **Задание 9**

В полученных файлах tnved\*\_utf.txt с помощью редактора nano заменить прямые разделители «|» на двойные «||» во всем тексте. Вывести содержимое на экран терминала. Переместить данные файлы далее в каталог /gold/files. Вывести итоговый результат работы в виде дерева каталога /my\_test\_folder

Запустим редактор **nano**, дальше введём сочетание клавиш ‘ctrl’ + ‘\’, далее вводим интересующий нас символ (‘|’) и символ на который хотим заменить (‘||’). Нажимаем A чтобы применить для всех.



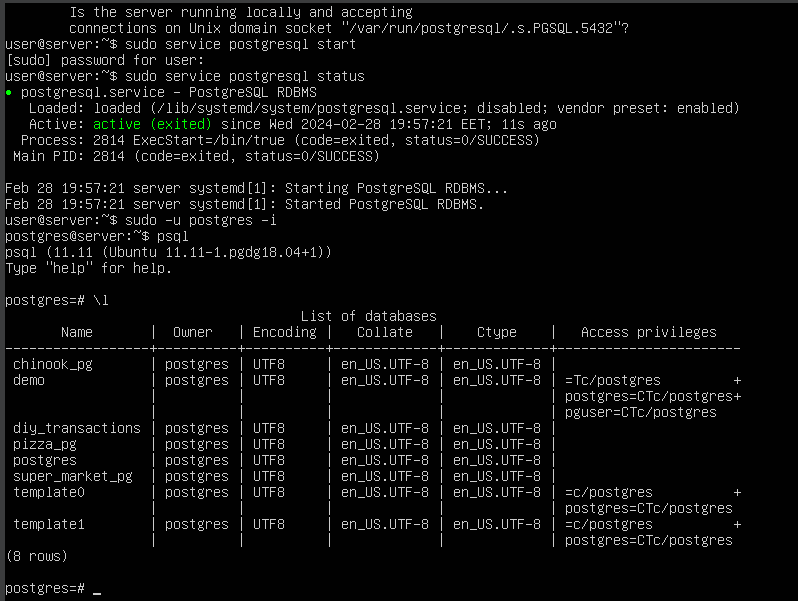
Ну и выведем дерево каталогов.



# **Задание 10**

Запустить PostgreSQL на виртуальной машине посредством соединения с командной строкой и ssh. Проверить наличие таблиц с данными в БД.

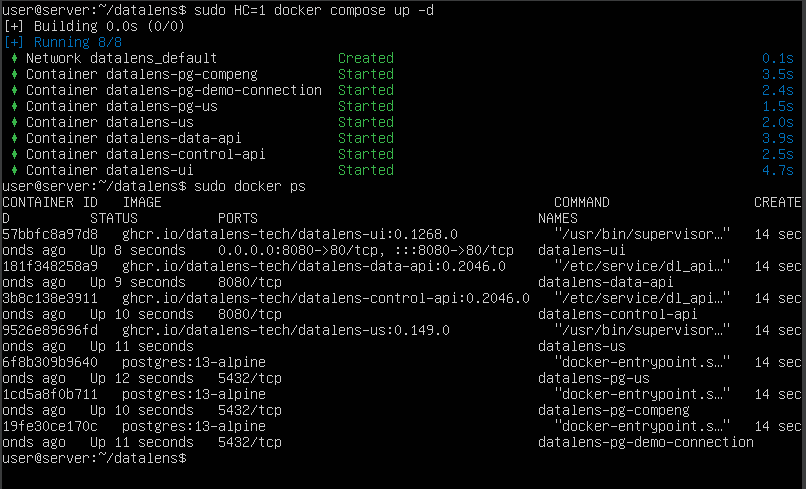
Запустим postgresql с помощью команды **sudo service postgresql start**, проверим что процесс активен. Проверим наличие таблиц с данными в БД.



# **Задание 11**

Запустить DataLens на виртуальной машине посредством соединения с командной строкой и ssh. Проверить работоспособность DataLens с помощью веб-браузера, открыв сайт DataLens

Перейдём в папку ~/datalens и запустим docker.



Теперь через веб-браузер проверим работоспособность DataLens.

