

```
!pip install -U scikit-learn

Requirement already satisfied: scikit-learn in
/usr/local/lib/python3.12/dist-packages (1.6.1)
Collecting scikit-learn
  Downloading scikit_learn-1.7.1-cp312-cp312-
manylinux2014_x86_64.manylinux_2_17_x86_64.whl.metadata (11 kB)
Requirement already satisfied: numpy>=1.22.0 in
/usr/local/lib/python3.12/dist-packages (from scikit-learn) (2.0.2)
Requirement already satisfied: scipy>=1.8.0 in
/usr/local/lib/python3.12/dist-packages (from scikit-learn) (1.16.1)
Requirement already satisfied: joblib>=1.2.0 in
/usr/local/lib/python3.12/dist-packages (from scikit-learn) (1.5.2)
Requirement already satisfied: threadpoolctl>=3.1.0 in
/usr/local/lib/python3.12/dist-packages (from scikit-learn) (3.6.0)
Downloading scikit_learn-1.7.1-cp312-cp312-
manylinux2014_x86_64.manylinux_2_17_x86_64.whl (9.5 MB)
  9.5/9.5 MB 68.6 MB/s eta
0:00:00
  ptning uninstall: scikit-learn
    Found existing installation: scikit-learn 1.6.1
    Uninstalling scikit-learn-1.6.1:
      Successfully uninstalled scikit-learn-1.6.1
Successfully installed scikit-learn-1.7.1

import pandas as pd
import numpy as np
pd.set_option('max_colwidth', None)
import matplotlib.pyplot as plt
import seaborn as sns
from scipy.spatial.distance import cdist, pdist
from sklearn.metrics import silhouette_score
import torch
from sentence_transformers import SentenceTransformer
from sklearn.cluster import KMeans
from sklearn.metrics import classification_report

from google.colab import drive
drive.mount('/content/drive')
reviews=pd.read_csv()
data=reviews.copy()
data.loc[1,'Text']

from google.colab import files
uploaded=files.upload()
data=pd.read_csv('/content/news_articles.csv')
data.head()
data = data.copy()
```

```
<IPython.core.display.HTML object>

Saving news_articles.csv to news_articles.csv

data.loc[1,'Text']

{"type": "string"}
```

data.head(5)

```
{"summary": "{\n    \"name\": \"data\", \n    \"rows\": 2225, \n    \"fields\": [\n        {\n            \"column\": \"Text\", \n            \"properties\": {\n                \"dtype\": \"string\", \n                \"num_unique_values\": 2127, \n                \"samples\": [\n                    \"Blair congratulates Bush on win\\n \\n Tony Blair has said he looks forward to continuing his strong relationship with George Bush and working with him during his second term as president.\\n \\n Mr Bush's re-election came at a crucial time for a world that was \\"fractured, divided and uncertain\\\"", Mr Blair said. It had to be brought together, he added, saying action was needed on poverty, the Middle East and the conditions on which terrorists prey. Mr Blair said states had to work with the US to fight global terrorism.\\n \\n But there was a need to recognise it would not be defeated by \\"military might alone but also by demonstrating the strength of our common values\\\"", he added. Solving the conflict in the Middle East was the world's single most \\"pressing political challenge\\\"", of the present day, Mr Blair warned. The prime minister also urged Europe and the US to \\"build anew their alliance\\\".\\n \\n \\"All of us in positions of leadership, not just President Bush, have a responsibility to rise to this challenge. It is urgent that we do so.\\\"", Mr Blair also paid tribute to Democrat John Kerry's campaign, saying he had helped make the presidential election \\"a true celebration of American democracy\\\". The election of the US president was significant for the world but particularly so for Britain because of its special relationship, he added. Earlier Tory leader Michael Howard sent Mr Bush his \\"warmest congratulations\\\"", saying: \\"We look to the president to be a unifying force for those all over the world who share our determination to defend freedom.\\\" Liberal Democrat leader Charles Kennedy welcomed the fact there had been a quick conclusion to the election, unlike in 2000.\\n \\n Mr Bush's first task was to \\"rebuild a sense of domestic purpose\\\" within the US, he said. Mr Kennedy said: \\"Internationally, it is to be hoped that a second term will see a more sensitive approach to relations with long-standing allies, not least for the global efforts to combat terrorism.\\\" Lib Dem foreign affairs spokesman Menzies Campbell said a win by Mr Kerry would have given Mr Blair the chance of a fresh start, adding it was almost as if there was an \\"umbilical cord\\\" between Mr Bush and the UK premier.\\n \\n \\"Europeans must hope that his administration will be much more multilateral in character, and that he will act swiftly to rebuild the Atlantic partnership which
```

is so vital to security. \\\"Iraq will remain an issue of potential division for some time to come.\\\" Even before the result became clear, Mr Blair was being urged to push for action on climate change at his first meeting with whichever candidate won.\\n \\n Liberal Democrat environment spokesman Norman Baker underlined the issue of global warming during a Commons debate on Anglo-American relations on Wednesday. Environment Secretary Margaret Beckett has said the US will act on global warming despite George Bush's refusal to sign up to the Kyoto protocol on carbon emissions. Public opinion would force change, she told BBC news. But Myron Ebell, an adviser on climate change to President Bush, has said there would be no change in the US stance and rejected the threat of climate change. He claimed the US was the only country with independent scientists.\\n\\n \\\"China Aviation seeks rescue deal\\n \\n Scandal-hit jet fuel supplier China Aviation Oil has offered to repay its creditors \$220m (\\u00c2\\u00a3117m) of the \$550m it lost on trading in oil futures.\\n \\n The firm said it hoped to pay \$100m now and another \$120m over eight years. With assets of \$200m and liabilities totalling \$648m, it needs creditors' backing for the offer to avoid going into bankruptcy. The trading scandal is the biggest to hit Singapore since the \$1.2bn collapse of Barings Bank in 1995. Chen Jiulin, chief executive of China Aviation Oil (CAO), was arrested by Changi Airport by Singapore police on 8 December. He was returning from China, where he had headed when CAO announced its trading debacle in late-November. The firm had been betting heavily on a fall in the price of oil during October, but prices rose sharply instead.\\n \\n Among the creditors whose backing CAO needs for its restructuring plan are banking giants such as Barclay's Capital and Sumitomo Mitsui, as well as South Korean firm SK Energy. Of the immediate payment, the firm - China's biggest jet fuel supplier - said it would be paying \$30m out of its own resources. The rest would come from its parent company, China Aviation Oil Holding Company in Beijing. The holding company, owned by the Chinese government, holds most of CAO's Singapore-listed shares. It cut its holding from 75% to 60% on 20 October.\\n\\n \\\"Yukos heading back to US courts\\n \\n Russian oil and gas company Yukos is due in a US court on Thursday as it continues to fight for its survival.\\n \\n The firm is in the process of being broken up by Russian authorities in order to pay a \$27bn (\\u00c2\\u00a314bn) tax bill. Yukos filed for

bankruptcy in the US, hoping to use international business law to halt the forced sale of its key oil production unit, Yuganskneftegas. The unit was however sold for \$9.4bn to state oil firm Rosneft but only after the state auction had been disrupted.\\n \\n Yukos lawyers now say the auction violated US bankruptcy law. The company and its main shareholders have vowed to go after any company that buys its assets, using all and every legal means. The company wants damages of \$20bn, claiming Yuganskneftegas was sold at less than market value. Judge Letitia Clark will hear different motions, including one from Deutsche Bank to throw out the Chapter 11 bankruptcy filing. The German lender is one of six banks that were barred from providing financing to

Gazprom, the Russian state-owned company that was expected to win the auction for Yuganskneftegas. Deutsche Bank, which is also an advisor to Gazprom, has called on the US court to overturn its decision to provide Yukos with bankruptcy protection.

Lifting the injunction would remove the uncertainty that surrounds the court case and clarify Deutsche Bank's business position, analysts said.

Analysts are not optimistic about Yukos' chances in court. Russian President Vladimir Putin and the country's legal authorities have repeatedly said that the US has no jurisdiction over Yukos and its legal wranglings. On top of that, the firm only has limited assets in the US. Yukos has won small victories, however, and is bullish about its chances in court.

"Do we have an ability to influence what happens? We think we do," said Mike Lake, a Yukos spokesman.

The litigation risks are real,

analyst Vadim Mitroshin The dispute with the Russian authorities is partly driven by President Putin's clampdown on the political ambitions of ex-Yukos boss Mikhail Khodorkovsky. Mr Khodorkovsky is in jail on charges of fraud and tax evasion.

"semantic_type": "Text",
"description": "Yukos bankruptcy 'not US matter'",
"type": "dataframe", "variable_name": "data"}

data.shape
(2225, 1)

data.tail(5)

{"summary": {"name": "data", "rows": 5, "fields": [{"column": "Text", "properties": {"dtype": "string", "num_unique_values": 5, "samples": ["Yukos bankruptcy 'not US matter']}}]}, "text": "Russian authorities do not have to abide by any US court decisions taken with regard to troubled oil giant Yukos, a Houston court has been told."}
Legal expert William Butler said there was no treaty between the US and Russia to recognise the other's legal rulings. That meant Moscow would not have to adhere to US rulings in the Yukos case.
Yukos says a US court was entitled to declare it bankrupt before its Yugansk unit was sold, since it has a US subsidiary and local bank accounts.
Yukos made its surprise Chapter 11 bankruptcy filing in Houston in December in an unsuccessful attempt to halt the auction of Yugansk, its main oil producing unit, by Russian authorities.
Yugansk was sold to help pay off a \$27.5bn (\u00c2\u00a314.5bn) back tax bill. It was bought for \$9.4bn by a previously unknown group, which was in turn bought by state-controlled oil company Rosneft. The US court's jurisdiction has been challenged by Deutsche Bank and Gazpromneft, a former unit of Russian gas monopoly Gazprom which is due to merge with Rosneft. Deutsche Bank maintains the case has no place in a US court because Yukos has no assets in the US, apart from two bank accounts and a house in Houston owned by its chief finance

officer Bruce Misamore. Deutsche Bank is involved in the case because it is itself being sued by Yukos. It had agreed to loan Gazpromneft the money to bid for Yugansk. US bankruptcy judge Letitia Clark, who issued an injunction in December to try and prevent the Yugansk sale, has said she will rule \"pretty promptly, however I do not anticipate ruling on it before next Tuesday\".
Yukos has claimed it sought help in the US because other forums - Russian courts and the European Court of Human Rights - were either unfriendly or offered less protection. It has claimed that Russia imposed the huge tax bill and forced the sale of Yugansk as part of a campaign to destroy Yukos and its former owner Mihkail Khodorkovsky, who is facing a 10-year prison term in Russia for fraud and tax evasion. Yukos' parent company, the Gibraltar-based Menatep Group, is suing Russia in Europe for \$28.3bn in financial damages. The company is also seeking \$20bn in a separate US lawsuit against Rosneft and Gazprom for their role in the sale of Yugansk.

"US trade gap hits record in 2004"
The gap between US exports and imports hit an all-time high of \$671.7bn (€3484bn) in 2004, latest figures show.

The Commerce Department said the trade deficit for all of last year was 24.4% above the previous record - 2003's imbalance of \$496.5bn. The deficit with China, up 30.5% at \$162bn, was the largest ever recorded with a single country. However, on a monthly basis the US trade gap narrowed by 4.9% in December to €356.4bn. The US consumer's appetite for all things from oil to imported cars, and even wine and cheese, reached record levels last year and the figures are likely to spark fresh criticism of President Bush's economic policies.

Democrats claim the administration has not done enough to clamp down on unfair foreign trade practices. For example, they believe China's currency policy - which US manufacturers claim has undervalued the yuan by as much as 40% - has given China's rapidly expanding economy an unfair advantage against US competitors.

Meanwhile, the Bush administration argues that the US deficit reflects the fact the America is growing at faster rate than the rest of the world, spurring on more demand for imported goods. Some economists say this may allow an upward revision of US economic growth in the fourth quarter. But others point out that the deficit has reached such astronomical proportions that foreigners may choose not to hold as many dollar-denominated assets, which may in turn harm growth. For all of 2004, US exports rose 12.3% to \$1.15 trillion, but imports rose even faster by 16.3% to a new record of \$1.76 trillion. Foreign oil exports surged by 35.7% to a record \$180.7bn, reflecting the rally in global oil prices and increasing domestic demand. Imports were not affected by the dollar's weakness last year.

"We expect the deficit to continue to widen in 2005 even if the dollar gets back to its downward trend," said economist Marie-Pierre Ripert at IXIS.

"Survey confirms property slowdown"
Government figures have confirmed a widely reported slowdown of the UK's housing market in late 2004.

House prices were 11.8% higher on the year in the last quarter of

2004, down from 16.3% in the July-to-September quarter, the Land Registry said. The average house price in England and Wales was \u00c2\u00a3182,920, down from \u00c2\u00a3187,971 in July-September. The volume of sales between October and December dropped by nearly a quarter from the same period in 2003. The government figures are the first official confirmation of falls in the market at the end of 2004. Land Registry figures are less up to date than those of banks and building societies, since they record completions not mortgage approvals. However, the figures are viewed as the most accurate measure of house prices as they include all property transactions, including cash sales.\n\nThe cost of buying a home fell in seven out of 10 regions between the third and fourth quarters of 2004.\n\nThe biggest annual gains were made in Wales, where house prices were up by 23% in the fourth quarter. House prices rose the slowest in Greater London, being up by 6%. In the capital, the volume of sales fell by 23% from 36,185 in 2003 to 28,041 for the same period in 2004. There was also a decline in the number of million-pound properties sold in the capital, with 436 properties over \u00a31m sold compared to 469 for the same period in 2003. Although the figures point to a slowdown in the market, the most recent surveys from Nationwide and Halifax have indicated the market may be undergoing a revival. After registering falls at the back end of 2004, Halifax said house prices rose by 0.8% in January and Nationwide reported a rise of 0.4% in the first month of the year. Members of the Bank of England's rate-setting committee will make their latest decision on interest rates on Thursday.\n\n],\n \"semantic_type\": \"/\",\\n \"description\": \"\\n }\\n }\\n]\\n },\n \"type\":\"dataframe\"}

```
data.duplicated().sum()
data=data.drop_duplicates()
data.reset_index(drop=True,inplace=True)
data.duplicated().sum()

np.int64(0)

data.shape
(2127, 1)

!pip install -U sentence-transformers -q

import warnings
warnings.filterwarnings("ignore")

data.loc[3, 'Text']

{"type":"string"}

data.isnull().sum()
```

```
Text      0
dtype: int64

data.duplicated().sum()

np.int64(0)

data.reset_index(drop=True, inplace=True)

data.duplicated().sum()

np.int64(0)

data.shape
(2127, 1)

!pip install hf_xet

Requirement already satisfied: hf_xet in
/usr/local/lib/python3.12/dist-packages (1.1.9)

from sentence_transformers import SentenceTransformer
#Defining the model
model = SentenceTransformer('sentence-transformers/all-MiniLM-L6-v2')

{"model_id": "5863f20704c049d887864fb77d34a24d", "version_major": 2, "version_minor": 0}

{"model_id": "9c342ebae5d64b24aa6b413910c166d3", "version_major": 2, "version_minor": 0}

 {"model_id": "1054c69a6775426487d31afd2644c812", "version_major": 2, "version_minor": 0}

 {"model_id": "3d4337acd16348c884d822fb601411ee", "version_major": 2, "version_minor": 0}

 {"model_id": "54dbe4e2abaa435ba8e4db91eab78eab", "version_major": 2, "version_minor": 0}

 {"model_id": "f11944d882e24df085d513eea7ff4eeb", "version_major": 2, "version_minor": 0}

 {"model_id": "b8f7da8fa8804aaeae73c8aefd28b5cf", "version_major": 2, "version_minor": 0}

 {"model_id": "bdaa55b52a5e4eea8ae5a3eb75caa58a", "version_major": 2, "version_minor": 0}

 {"model_id": "aa4d808fbcd49a7ba1d8e5145b5a397", "version_major": 2, "version_minor": 0}
```

```
{"model_id": "0bcc8b098a084c21ae864129a737b07b", "version_major": 2, "version_minor": 0}

{"model_id": "a14efbeaec9b45ae9dd94cb4e116554b", "version_major": 2, "version_minor": 0}

device = torch.device("cuda" if torch.cuda.is_available() else "cpu")
embedding_matrix = model.encode(data['Text'], show_progress_bar=True,
device=device)
embedding_matrix.shape

{"model_id": "7871328dd3364ad3a8c5f7a1b44f95d7", "version_major": 2, "version_minor": 0}

(2127, 384)

embedding_matrix = model.encode(data['Text'], show_progress_bar=True,
device=device)

{"model_id": "324f0ff843e94880986c7eed119e3593", "version_major": 2, "version_minor": 0}

embedding_matrix.shape

(2127, 384)

len(embedding_matrix[0])

384

a= "i love apple"
b= "apple is a fruit"
c= "i like this table"

def cosine_score(text1,text2):
    # encoding the text
    embeddings1 = model.encode(text1)
    embeddings2 = model.encode(text2)
    # calculating the L2 norm of the embedding vector
    norm1 = np.linalg.norm(embeddings1)
    norm2 = np.linalg.norm(embeddings2)
    # computing the cosine similarity
    cosine_similarity_score =
((np.dot(embeddings1,embeddings2))/(norm1*norm2))
    return cosine_similarity_score

print(cosine_score(a,b))
print(cosine_score(b,c))
print(cosine_score(a,c))
```

```

0.55750245
0.07198238
0.27396193

from sentence_transformers import util
embeddings1 = model.encode(a)
embeddings2 = model.encode(b)
embeddings3 = model.encode(c)
print(util.cos_sim(embeddings1, embeddings2))
print(util.cos_sim(embeddings2, embeddings3))
print(util.cos_sim(embeddings1, embeddings3))

tensor([[0.5575]])
tensor([[0.0720]])
tensor([[0.2740]])

def top_k_similar_sentences(embedding_matrix, query_text, k):
    query_embedding = model.encode(query_text)
    score_vector = np.dot(embedding_matrix, query_embedding)
    top_k_indices = np.argsort(score_vector)[-1][:-1]
    return data.loc[list(top_k_indices), 'Text']

query_text = "Budget for elections"
# displaying the top 3 similar sentences
top_k_reviews = top_k_similar_sentences(embedding_matrix, query_text,
3)

for i in top_k_reviews:
    print(i, end="\n")

print("*****")
print("\n")

```

'Debate needed' on donations cap

A cap on donations to political parties should not be introduced yet, the elections watchdog has said.

Fears that big donors can buy political favours have sparked calls for a limit. In a new report, the Electoral Commission says it is worth debating a £10,000 cap for the future but now is not the right time to introduce it. It also says there should be more state funding for political parties and candidates should be able to spend more on election campaigning.

There were almost £68m in reported donations to political parties in 2001, 2002 and 2003, with nearly £12m of them from individual gifts worth more than £1m. The rules have already been changed so the public can see who gives how much to the parties but the report says

there are still public suspicions. The commission says capping donations would mean taxpayers giving parties more cash - something which would first have to be acceptable to the public and shown to work. "While we are not in principle opposed to the introduction of a donation cap, we do not believe that such a major departure from the existing system now would be sensible," says its report. If there was to be a cap, it should be Â£10,000 - a small enough amount to make a difference but which would have banned Â£56m in donations between 2001 and 2003.

Even without changes the commission does urge political parties to seek out more small-scale donations and suggests there should be income tax relief for gifts under Â£200. It also suggests increasing state funding for parties to Â£3m so help can be extended to all parties with at least two members in the House of Commons, European Parliament, Scottish Parliament, Welsh Assembly or Northern Ireland Assembly. And it suggests new ways of boosting election campaigning, seen as a way of improving voter turnout. All local election candidates should be entitled to a free mailshot for campaign leaflets, says the watchdog. And there should be a shift in the amount of money allowed to be spent at elections from a national level to a local level to help politicians engage better with voters.

The report suggests doubling the money which can be spent by candidates, while cutting national spending limits from Â£20m to Â£15m. The commission also says the spending limits for general elections should cover the four months before the poll - as happens with other elections. Electoral Commission chairman Sam Younger said: "There is no doubt that political parties have a vital role to play in maintaining the health of our democracy and for this they need to be adequately resourced. "Our research has shown that people want to be more informed about party politics and that they want politicians to be more visible and accessible. "The public are reluctant for the state to fund parties but at the same time are unhappy with large private donations." He called for a wider public debate on party funding to find the consensus needed for radical changes to the current system.

Bush budget seeks deep cutbacks

President Bush has presented his 2006 budget, cutting domestic spending in a bid to lower a record deficit projected to peak at \$427bn (Â£230bn) this year.

The \$2.58 trillion (Â£1.38 trillion) budget submitted to Congress affects 150 domestic programmes from farming to the environment, education and health. But foreign aid is due to rise by 10%, with more

money to treat HIV/Aids and reward economic and political reform. Military spending is also set to rise by 4.8%, to reach \$419.3bn. The budget does not include the cost of running military operations in Iraq and Afghanistan, for which the administration is expected to seek an extra \$80bn from Congress later this year.

Congress will spend several months debating George W Bush's proposal.

The state department's planned budget would rise to just under \$23bn - a fraction of the defence department's request - including almost \$6bn to assist US allies in the "war on terror". However, the administration is keen to highlight its global effort to tackle HIV/Aids, the BBC's Jonathan Beale reports, and planned spending would almost double to \$3bn, with much of that money going to African nations. Mr Bush also wants to increase the amount given to poorer countries through his Millennium Challenge Corporation. The scheme has been set up to reward developing countries that embrace what the US considers to be good governance and sound policies. Yet Mr Bush's proposed spending of \$3bn on that project is well below his initial promise of \$5bn.

A key spending line missing from proposals is the cost of funding the administration's proposed radical overhaul of Social Security, the pensions programme on which many Americans rely for their retirement income.

Some experts believe this could require borrowing of up to \$4.5 trillion over a 20-year period. Neither does the budget include any cash to purchase crude oil for the US emergency petroleum stockpile. Concern over the level of the reserve, created in 1970s, has led to rises in oil prices over the past year. The Bush administration will instead continue to fill the reserve by taking oil - rather than cash - from energy companies that drill under federal leases.

The outline proposes reductions in budgets at 12 out of 23 government agencies including cuts of 9.6% at Agriculture and 5.6% at the Environmental Protection Agency. The spending plan for the year beginning 1 October is banking on a healthy US economy to boost government income by 6.1% to \$2.18 trillion. Spending is forecast to grow by 3.5% to \$2.57 trillion. But the budget is still the tightest yet under Mr Bush's presidency. "In order to sustain our economic expansion, we must continue pro-growth policies and enforce even greater spending restraint across federal government," Mr Bush said in his budget message to Congress. Mr Bush has promised to halve the US's massive budget deficit within five years. The deficit, partly the result of massive tax cuts early in Mr Bush's presidency, has been a key factor in pushing the US dollar lower. The independent Congressional Budget Office estimates that the shortfall could shrink to little more than \$200bn by 2009, returning to the surpluses seen in

the late 1990s by 2012. But its estimates depend on the tax cuts not being made permanent, in line with the promise when they were passed that they would "sunset", or disappear, in 2010. Most Republicans, however, want them to stay in place. And the figures also rely on the "Social Security trust fund" - the money set aside to cover the swelling costs of retirement pensions - being offset against the main budget deficit.

Will the Budget bring out smiling voters?

As Tory spokesman Oliver Letwin said - any chancellor would use his pre-election budget to offer some vote winning sweeteners, wouldn't he.

And everyone does, indeed, expect Gordon Brown to do just that in his last budget before the expected polling day of 5 May. There will be plenty of talk about taking no risks with the economy or handing out irresponsible giveaways. But Mr Brown will stun Westminster and break just about every historical precedent if he fails to do something designed to put a smile on voters' faces and make them more inclined to back Labour in the election. And there has already been speculation about possible tax reductions for the poorest and increasing the threshold on stamp duty in this week's Budget. The aim of his package will be to keep any disillusioned core Labour voters in the fold, while ensuring the middle England voters who gave Tony Blair two election victories don't desert him at the third poll.

And, needless to say, there will be plenty of analysis of what impact the Budget will have on Mr Brown's own ambitions to replace Mr Blair as prime minister at some point after a third win. But there is a shadow hanging over this pre-election performance - in the shape of £11 billion or thereabouts. That is the size of the financial "black hole" the Tories, backed by some independent forecasters, believe Mr Brown will have to fill with tax increases after the election.

The opening shot in that battle was fired at the weekend with claims a Treasury leak suggested Mr Brown was ready to slap capital gains tax on home sales to raise some of that cash. It was immediately denied, but the Tories remain suspicious, claiming that, as Mr Letwin said, if the money does not come from there, where will it come from.

Mr Brown will undoubtedly claim the forecasts are simply wrong and that he will have absolutely no need to raise taxes after the election, should Labour win. Previous gloomy forecasts proved wrong, he will argue, while suggesting that only by sticking with him can Britain continue to have a sound economy, low unemployment and high public spending. The opposition will reject that by claiming they can

keep the economy sound, increase spending and cut taxes at the same time.

The Liberal Democrats will also promise to run a sound economy, but based on increased taxes to fund spending and, amongst other things, abolishing the council tax in favour of a local income tax. And there will be prolonged argument over which of the parties can make the greatest savings in Whitehall and beyond to fund their policies. But probably what that boils down to is a simple question of which of the parties the voters most trust to keep the economy stable and avoid any unpleasant surprises in the coming years. And it is probably still the answer to that one question that will overwhelmingly decide the outcome of the general election.

```
query_text = "High imports and exports"
# displaying the top 3 similar sentences
top_k_reviews = top_k_similar_sentences(embedding_matrix, query_text,
3)

for i in top_k_reviews:
    print(i, end="\n")

print("*****")
print("\n")
```

US trade gap ballooned in October

The US trade deficit widened by more than expected in October, hitting record levels after higher oil prices raised import costs, figures have shown

The trade shortfall was \$55.5bn (£29bn), up 9% from September, the Commerce Department said. That pushed the 10 month deficit to \$500.5bn. Imports rose by 3.4%, while exports increased by only 0.6%. A weaker dollar also increased the cost of imports, though this should help drive export demand in coming months. "Things are getting worse, but that's to be expected," said David Wyss of Standard & Poor's in New York. "The first thing is that when the dollar goes down, it increases the price of imports. "We are seeing improved export orders. Things seem to be going in the right direction."

Despite this optimism, significant concerns remain as to how the US will fund its trade and budget deficits should they continue to widen. Another problem highlighted by analysts was the growing trade gap with China, which has been accused of keeping its currency artificially

weak in order to boost exports. The US imported almost \$20bn worth of goods from China during October, exporting a little under \$3bn. "It seems the key worry that has existed in the currency market still remains," said Anthony Crescenzi, a bond strategist at Miller Tabak in New York. The trade deficit and the shortfall with China "are big issues going forward". The Commerce Department figures caused the dollar to weaken further despite widespread expectations that the Federal Reserve will raise interest rates for a fifth time this year. Borrowing costs are tipped to rise by a quarter of a percentage point to 2.25% at a Fed meeting later on Tuesday.

Trade gap narrows as exports rise

The UK's trade gap narrowed in November, helped by a 7.5% rise in exports outside the European Union.

According to the Office for National Statistics, the difference between what the UK exported and imported was £3.1bn (\$5.8bn), down from October's £3.6bn. Overall UK exports - including both goods and services - rose by more than 3.2% to £24.8bn, although total imports rose again to a new record of £27.9bn. The deficit for goods alone was £4.6bn, down from October's £5bn. During November the UK exported £16.9bn worth of goods, but imported £21.5bn. The cumulative deficit for the first eleven months of 2004 now stands at £36.3bn, £4.5bn higher than the same period in 2003.

November saw an improvement in export levels to both the European Union and the rest of the world, the Office for National Statistics (ONS) said. EU exports rose 2%, fuelled by an increase in sales of chemicals. Non-EU exports shot up 7.5%, with growth seen across a range of manufacturing sectors including cars, consumer durables and chemicals. The export boost offset a 1% rise in imports. Non-EU imports rose 3%, but the growth in goods entering the UK from the EU slowed to 0.5%.

The UK's deficit with the EU fell to £1.9bn from £2.1bn, while its non-EU shortfall dropped to £2.7bn from £2.9bn in October.

The country's surplus on trade-in-services remained steady at £1.5bn for the fifth month in a row. Paul Dales, UK economist for Capital Economics, said the figures represented an improvement on recent months. However, he stressed that the long-term prognosis for exports was still uncertain. "The figures are a lot better than expected but the trend still remains poor," he said. "There have been some very encouraging signs that the UK export recovery is starting to take hold. But there is a danger that this could be held back by the ongoing weakness of domestic demand on the continent."

US trade gap hits record in 2004

The gap between US exports and imports hit an all-time high of \$671.7bn (£484bn) in 2004, latest figures show.

The Commerce Department said the trade deficit for all of last year was 24.4% above the previous record - 2003's imbalance of \$496.5bn. The deficit with China, up 30.5% at \$162bn, was the largest ever recorded with a single country. However, on a monthly basis the US trade gap narrowed by 4.9% in December to £56.4bn. The US consumer's appetite for all things from oil to imported cars, and even wine and cheese, reached record levels last year and the figures are likely to spark fresh criticism of President Bush's economic policies.

Democrats claim the administration has not done enough to clamp down on unfair foreign trade practices. For example, they believe China's currency policy - which US manufacturers claim has undervalued the yuan by as much as 40% - has given China's rapidly expanding economy an unfair advantage against US competitors.

Meanwhile, the Bush administration argues that the US deficit reflects the fact the America is growing at faster rate than the rest of the world, spurring on more demand for imported goods. Some economists say this may allow an upward revision of US economic growth in the fourth quarter. But others point out that the deficit has reached such astronomical proportions that foreigners many choose not to hold as many dollar-denominated assets, which may in turn harm growth. For all of 2004, US exports rose 12.3% to \$1.15 trillion, but imports rose even faster by 16.3% to a new record of \$1.76 trillion. Foreign oil exports surged by 35.7% to a record \$180.7bn, reflecting the rally in global oil prices and increasing domestic demand. Imports were not affected by the dollar's weakness last year. "We expect the deficit to continue to widen in 2005 even if the dollar gets back to its downward trend," said economist Marie-Pierre Ripert at IXIS.

```
meanDistortions = []
clusters = range(2, 11)
for k in clusters:
    clusterer = KMeans(n_clusters=k, random_state=1)
    clusterer.fit(embedding_matrix)
    prediction = clusterer.predict(embedding_matrix)
    distortion = sum(
```

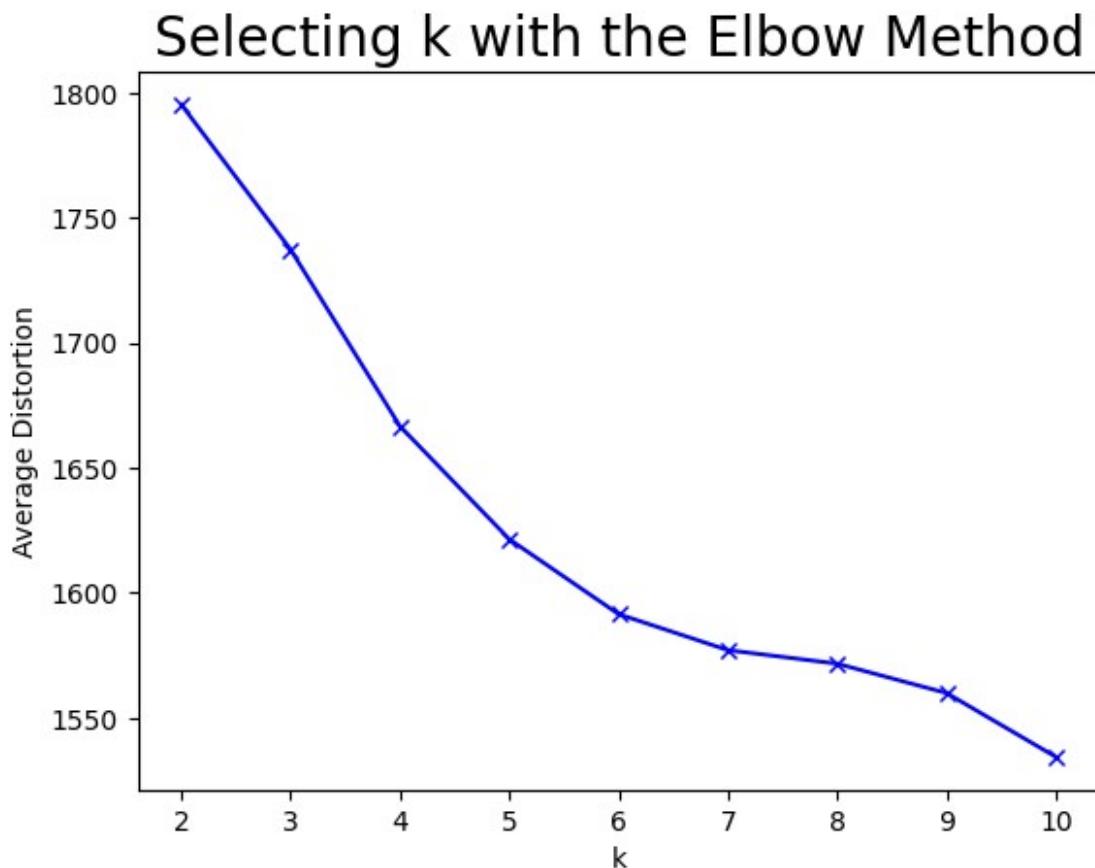
```

        np.min(cdist(embedding_matrix, clusterer.cluster_centers_,
"euclidean"), axis=1) ** 2
    )
meanDistortions.append(distortion)
print("Number of Clusters:", k, "\tAverage Distortion:",
distortion)

Number of Clusters: 2      Average Distortion: 1795.3091146747045
Number of Clusters: 3      Average Distortion: 1737.2980165191386
Number of Clusters: 4      Average Distortion: 1666.670280185686
Number of Clusters: 5      Average Distortion: 1621.4823707705561
Number of Clusters: 6      Average Distortion: 1591.5027285873807
Number of Clusters: 7      Average Distortion: 1577.105400570846
Number of Clusters: 8      Average Distortion: 1571.684353511434
Number of Clusters: 9      Average Distortion: 1559.9432993656876
Number of Clusters: 10     Average Distortion: 1534.399210400378

plt.plot(clusters, meanDistortions, "bx-")
plt.xlabel("k")
plt.ylabel("Average Distortion")
plt.title("Selecting k with the Elbow Method", fontsize=20)
plt.show()

```



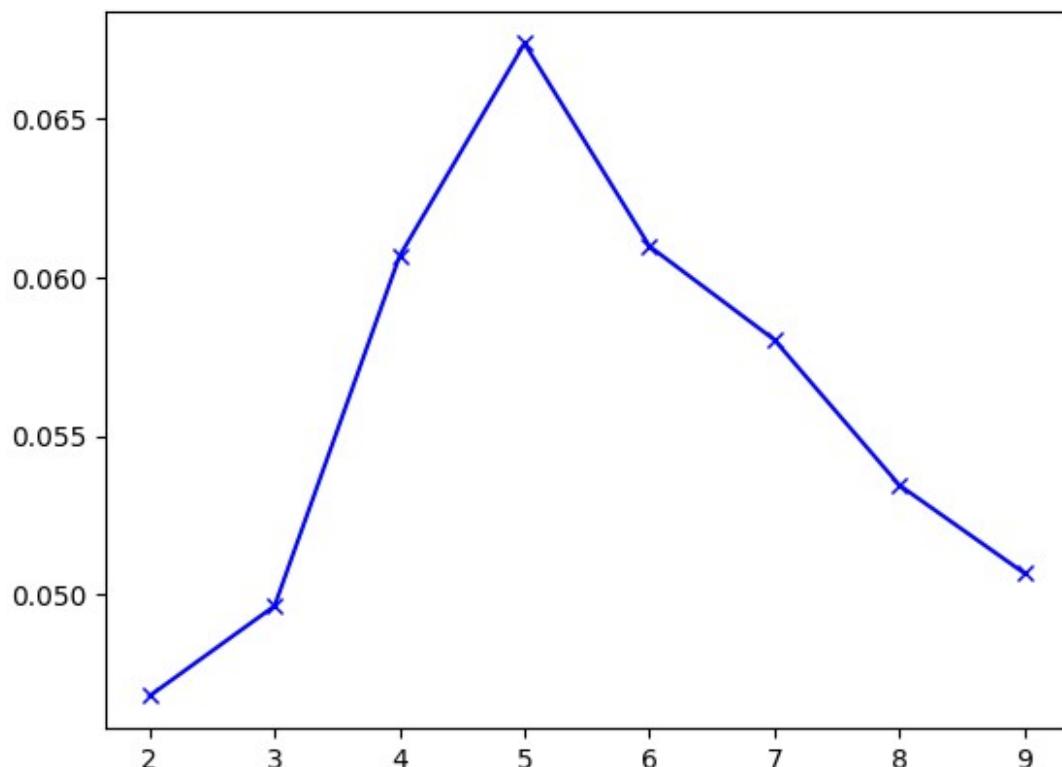
```

sil_score = []
cluster_list = range(2, 10)
for n_clusters in cluster_list:
    clusterer = KMeans(n_clusters=n_clusters, random_state=1)
    preds = clusterer.fit_predict(embedding_matrix)
    score = silhouette_score(embedding_matrix, preds)
    sil_score.append(score)
    print("For n_clusters = {}, the silhouette score is {}".format(n_clusters, score))

For n_clusters = 2, the silhouette score is 0.046811509877443314)
For n_clusters = 3, the silhouette score is 0.04964283108711243)
For n_clusters = 4, the silhouette score is 0.06065066531300545)
For n_clusters = 5, the silhouette score is 0.06735046207904816)
For n_clusters = 6, the silhouette score is 0.06096599996089935)
For n_clusters = 7, the silhouette score is 0.05802200734615326)
For n_clusters = 8, the silhouette score is 0.05344569683074951)
For n_clusters = 9, the silhouette score is 0.05067500099539757)

plt.plot(cluster_list, sil_score, "bx-")
plt.show()

```



```

n_categories = 5
kmeans = KMeans(n_clusters=n_categories,
random_state=1).fit(embedding_matrix)

```

```
centers = kmeans.cluster_centers_
centers

array([[-0.01785407,  0.00624583, -0.00339485, ..., -0.06009004,
       -0.03252158,  0.01114389],
       [ 0.00549149, -0.02162499,  0.01707268, ..., -0.03975475,
       -0.01703279,  0.01147507],
       [-0.01044312, -0.0430865 , -0.01626261, ..., -0.06730248,
       -0.0080532 ,  0.00465481],
       [-0.01394476, -0.03756041,  0.00739344, ..., -0.10312265,
       -0.01400104,  0.03134287],
       [-0.02824864, -0.03118991,  0.00693417, ..., -0.05258856,
       0.00538871,  0.03337956]], dtype=float32)

clustered_data = data.copy()
clustered_data['Category'] = kmeans.labels_
clustered_data.head()

{"summary": {"name": "clustered_data", "rows": 2127, "fields": [{"column": "Text", "dtype": "string"}, {"samples": ["Blair congratulates Bush on win\\n \\n Tony Blair has said he looks forward to continuing his strong relationship with George Bush and working with him during his second term as president.\\n \\n Mr Bush's re-election came at a crucial time for a world that was \\"fractured, divided and uncertain\\\"", Mr Blair said. It had to be brought together, he added, saying action was needed on poverty, the Middle East and the conditions on which terrorists prey. Mr Blair said states had to work with the US to fight global terrorism.\\n \\n But there was a need to recognise it would not be defeated by \\"military might alone but also by demonstrating the strength of our common values\\\"", he added. Solving the conflict in the Middle East was the world's single most \\"pressing political challenge\\\" of the present day, Mr Blair warned. The prime minister also urged Europe and the US to \\"build anew their alliance\\\".\\n \\"All of us in positions of leadership, not just President Bush, have a responsibility to rise to this challenge. It is urgent that we do so.\\\" Mr Blair also paid tribute to Democrat John Kerry's campaign, saying he had helped make the presidential election \\"a true celebration of American democracy\\\". The election of the US president was significant for the world but particularly so for Britain because of its special relationship, he added. Earlier Tory leader Michael Howard sent Mr Bush his \\"warmest congratulations\\\"", saying: \\"We look to the president to be a unifying force for those all over the world who share our determination to defend freedom.\\\" Liberal Democrat leader Charles Kennedy welcomed the fact there had been a quick conclusion to the election, unlike in 2000.\\n \\n Mr Bush's first task was to \\"rebuild a sense of domestic purpose\\\" within the US, he said. Mr Kennedy said: \\"Internationally, it is to be hoped that a second"]}], "properties": {"num_unique_values": 2127}}}
```

term will see a more sensitive approach to relations with long-standing allies, not least for the global efforts to combat terrorism." Lib Dem foreign affairs spokesman Menzies Campbell said a win by Mr Kerry would have given Mr Blair the chance of a fresh start, adding it was almost as if there was an "umbilical cord" between Mr Bush and the UK premier.
Europeans must hope that his administration will be much more multilateral in character, and that he will act swiftly to rebuild the Atlantic partnership which is so vital to security. "Iraq will remain an issue of potential division for some time to come." Even before the result became clear, Mr Blair was being urged to push for action on climate change at his first meeting with whichever candidate won.
Liberal Democrat environment spokesman Norman Baker underlined the issue of global warming during a Commons debate on Anglo-American relations on Wednesday. Environment Secretary Margaret Beckett has said the US will act on global warming despite George Bush's refusal to sign up to the Kyoto protocol on carbon emissions. Public opinion would force change, she told BBC news. But Myron Ebell, an adviser on climate change to President Bush, has said there would be no change in the US stance and rejected the threat of climate change. He claimed the US was the only country with independent scientists.

"China Aviation seeks rescue deal"
Scandal-hit jet fuel supplier China Aviation Oil has offered to repay its creditors \$220m (€220m) of the \$550m it lost on trading in oil futures.
The firm said it hoped to pay \$100m now and another \$120m over eight years. With assets of \$200m and liabilities totalling \$648m, it needs creditors' backing for the offer to avoid going into bankruptcy. The trading scandal is the biggest to hit Singapore since the \$1.2bn collapse of Barings Bank in 1995. Chen Jiulin, chief executive of China Aviation Oil (CAO), was arrested by Singapore police on 8 December. He was returning from China, where he had headed when CAO announced its trading debacle in late-November. The firm had been betting heavily on a fall in the price of oil during October, but prices rose sharply instead.

Among the creditors whose backing CAO needs for its restructuring plan are banking giants such as Barclay's Capital and Sumitomo Mitsui, as well as South Korean firm SK Energy. Of the immediate payment, the firm - China's biggest jet fuel supplier - said it would be paying \$30m out of its own resources. The rest would come from its parent company, China Aviation Oil Holding Company in Beijing. The holding company, owned by the Chinese government, holds most of CAO's Singapore-listed shares. It cut its holding from 75% to 60% on 20 October.

"Yukos heading back to US courts"
Russian oil and gas company Yukos is due in a US court on Thursday as it continues to fight for its survival.
The firm is in the process of being broken up by Russian authorities in order to pay a \$27bn (€27bn) tax bill. Yukos filed for bankruptcy in the US, hoping to use international business law to halt the forced sale of its key oil production unit, Yuganskneftegas. The unit was however sold for \$9.4bn to state oil firm Rosneft but only

after the state auction had been disrupted.\n\n Yukos lawyers now say the auction violated US bankruptcy law. The company and its main shareholders have vowed to go after any company that buys its assets, using all and every legal means. The company wants damages of \$20bn, claiming Yuganskneftegas was sold at less than market value. Judge Letitia Clark will hear different motions, including one from Deutsche Bank to throw out the Chapter 11 bankruptcy filing. The German lender is one of six banks that were barred from providing financing to Gazprom, the Russian state-owned company that was expected to win the auction for Yuganskneftegas. Deutsche Bank, which is also an advisor to Gazprom, has called on the US court to overturn its decision to provide Yukos with bankruptcy protection.\n\n Lifting the injunction would remove the uncertainty that surrounds the court case and clarify Deutsche Bank's business position, analysts said.\n\n Analysts are not optimistic about Yukos' chances in court. Russian President Vladimir Putin and the country's legal authorities have repeatedly said that the US has no jurisdiction over Yukos and its legal wranglings. On top of that, the firm only has limited assets in the US. Yukos has won small victories, however, and is bullish about its chances in court. \"Do we have an ability to influence what happens? We think we do,\"\" said Mike Lake, a Yukos spokesman.

\"The litigation risks are real,\"\" said Credit Suisse First Boston analyst Vadim Mitroshin. The dispute with the Russian authorities is partly driven by President Putin's clampdown on the political ambitions of ex-Yukos boss Mikhail Khodorkovsky. Mr Khodorkovsky is in jail on charges of fraud and tax evasion.\n\n],\n\n \"semantic_type\": \"\", \n\n \"description\": \"\"\n\n },\n\n {\n\n \"column\": \"Category\", \n\n \"properties\": {\n\n \"dtype\": \"int32\", \n\n \"num_unique_values\": 5,\n\n \"samples\": [\n\n 2, \n\n 0, \n\n 4\n\n],\n\n \"semantic_type\": \"\", \n\n \"description\": \"\"\n\n }\n\n }\n\n]\n\n }","type":"dataframe","variable_name":"clustered_data"}\n\nfor i in range(5):\n print("CLUSTER",i)\n print(clustered_data.loc[clustered_data.Category == i,\n'Text'].sample(5, random_state=1).values)\n\nprint("*****\n**")\n print("\n")\n\nCLUSTER 0\n["O'Sullivan keeps his powder dry\n\n When you are gunning for glory and ultimate success keeping the gunpowder dry is essential.\n\n Ireland coach Eddie O'Sullivan appears to have done that quite successfully in the run-up to this season's Six Nations Championship. He decreed after the 2003 World Cup that players should have a decent conditioning period during the year. That became a reality at the end of last summer with a 10-week period at the start of the this season."]

It may have annoyed his Scottish, and in particularly Welsh, cousins who huffed and puffed at the disrespect apparently shown to the Celtic League. We will say nothing of Mike Ruddock ''poaching'' eight of the Dragons side that faced Leinster on Sunday. But, like O'Sullivan, he was well within his rights, particularly when you are talking about the national side and pride that goes along with it. The IRFU has thrown their weight behind O'Sullivan, who must be glad that in the main, there is centrally-controlled contracts. Bar Keith Gleeson who is just returning from a broken leg, everyone of O'Sullivan's squad is fit, fresh and standing at the oche ready to launch this season's campaign.\n \n But I doubt whether O'Sullivan is going to gloat about the handling of his players. He is not that sort of person. However, he may look at the overworked and injury-hit England, Wales and France squads whose players have been overworked, and then pat himself on the back for his foresight. But there is still the question of turning up and transferring that freshness into positive results when the referee signals the start of the game. Already Ireland are being earmarked as hot favourites in many quarters to go the whole hog this season. A first Grand Slam since Karl Mullen's led the team to a clean sweep in 1948. With England and France visiting Lansdowne Road for the last time before the old darling is pulled down, everything looks perfectly placed. But in the days of yore that frightened the life out of any Irishman.\n \n Under the burden of great expectations, Ireland have crumpled. Take the Triple Crown-winning side of 1985 under Mick Doyle. They were expected to up the ante further for a Grand Slam, only the second in Ireland's history. What happened in 1986? Whitewashed. You see, Ireland, in any sport, love to be downsized. Then they can go out and prove a point to the contrary. It is the nature of the beast. But O'Sullivan's side are very capable of proving a salient point this season. After their first Triple Crown for 19 years, they can live up to their success and take a further step up the ladder. O'Sullivan has kept faith and displayed loyalty to his players, and they have repaid him in spades ... and there is more to come. He has some old dogs in his squad, but he will come to this season's championship with a different box of tricks, and a new verve to succeed. Ireland can indeed succeed, but just whisper it.\n"

'Roddick in talks over new coach\n \n Andy Roddick is reportedly close to confirming US Davis Cup assistant Dean Goldfine as his new coach.\n \n Roddick ended his 18-month partnership with Brad Gilbert on Monday, and Goldfine admits talks have taken place. "We had a really good conversation and we're on the same page in terms of what I expect from a player in commitment and what he wants," said Goldfine. "The reading I got from him is that I would have a lot of the qualities he's looking for in a coach." Speaking to told South Florida's Sun-Sentinel newspaper, Goldfine added: "That being said, from his standpoint, which is smart, he wants to cover all his bases. "I think Andy wants a long-term relationship and wants to make sure it's the right fit... the best fit." Goldfine, 39, has worked with Todd Martin and Roddick's close friend Mardy Fish, and was an

assistant coach with the US Olympic team. Martin is the other name to have been linked to the vacant post alongside Roddick.\n'

'Bryan twins keep US hopes alive\n \n The United States kept the Davis Cup final alive with victory in Saturday's doubles rubber, leaving Spain 2-1 ahead going into the final day.\n \n Masters Cup champions Mike and Bob Bryan thrashed Juan Carlos Ferrero and Tommy Robredo 6-0 6-3 6-2 in front of a partisan crowd in Seville. Victory would have given Spain the title but they were outclassed. In Sunday's reverse singles, Carlos Moya takes on Andy Roddick before Rafael Nadal faces Mardy Fish. "It feels good, but it's not going to be as good if we don't win two tomorrow," said Mike Bryan. "It feels good to give those guys another shot, and Spain has to go to sleep on that." Bob Bryan added: "I'm really confident in Andy winning that first match, and then anything can happen." Spain coach Jordi Arrese chose to rest 18-year-old Nadal in the doubles after his epic singles win over Roddick on Friday. He was replaced by former world number one Ferrero, but the Spanish pair were out of their depth against one of the world's best doubles teams.\n \n The 26-year-old Bryan twins have won all four of their Davis Cup matches this year. And they quickly silenced the huge crowd at the Olympic Stadium, racing through the opening set to love. The Spaniards then twice surrendered breaks of serve at the start of the second before the Bryans broke to go 5-3 ahead and served out. When Robredo dropped serve in the opening game of the third set the match was all but over, and the unflappable Bryan brothers powered on to an impressive win. Ferrero, who was upset to be dropped for Friday's singles, hinted at further dissatisfaction after the defeat. "It was a difficult game against the best doubles players," he said. "They have everything calculated and we had very little to do. "I was a bit surprised that I was named to play the doubles match because I hardly play doubles." Arrese said: "Juan Carlos hasn't played at all badly. He played the right way but the Bryans are great doubles players."\n'

'England coach faces rap after row\n \n England coach Andy Robinson is facing disciplinary action after criticising referee Jonathan Kaplan in his side's Six Nations defeat to Ireland.\n \n The Rugby Football Union (RFU) will investigate Robinson after deciding not to lodge a complaint against Kaplan. Robinson may even have to apologise for his comments in order to avoid sanction from the International Rugby Board. Robinson had said he was "livid" about Kaplan's decisions on Saturday to disallow two England "tries." The England coach went on to claim that "only one side was refereed".\n \n After reviewing tapes of the match, the RFU decided not to formally complain to the IRB over the standard of Kaplan's refereeing. Instead the RFU said in a statement they would, "set out any concerns the England team management may have in a confidential manner".\n \n An IRB spokesman said on the matter: "We take all breaches of the code very seriously. "Should the RFU resolve the issue to our satisfaction, as happened last month when the Scotland coach Matt Williams apologised for remarks made, it would be the end of the matter."\n \n Kaplan has

vigorously defended his performance in England's 19-13 defeat at Lansdowne Road and admitted he was "very disappointed" with Robinson's remarks. And the South African has been appointed to take charge of Scotland's match against Wales on 13 March.\n \n The RFU recently fined Northampton coach Budge Pountney £2,000 and imposed a six-week ban for his criticism of referee Steve Lander after a Premiership match.\n'

'England \'to launch ref protest'\n \n England will protest to the International Rugby Board (IRB) about the referee's performance in the defeat by Ireland, reports the Daily Mail.\n \n England coach Andy Robinson has called on ex-international referees Colin High and Steve Lander to analyse several of Jonathan Kaplan's decisions. "I want to go through the tape with Colin and Steve," Robinson told the Daily Mail. "I want to speak to the IRB about it. I think only one side was refereed." High, the Rugby Football Union's referees' manager, claimed Kaplan made three major errors which changed the outcome of Sunday's match. England were beaten 19-13 by the Irish in Dublin, their third straight defeat in the 2005 Six Nations. "The International Rugby Board will be disappointed," High told the Daily Mail. "Jonathan Kaplan is in the top 20 in the world but that wasn't an international performance. "It would not have been acceptable in the Zurich Premiership. "If one of my referees had done that, I would have had my backside kicked for making the appointment. "If any English referee refereed like that in a European match, there would be an inquest. No question about that. "If someone had performed like that, he would have been pulled from the next game."\n']

CLUSTER 1

['Malik rejects all-black MP lists\n \n A call for ethnic minority shortlists to boost the number of black and Asian MPs has been rejected by one of Labour's most senior Asians.\n \n Shahid Malik, who is on Labour's ruling NEC, accepted people's frustration but said there should be targets not lists to boost representation of minorities. Just 13 of Britain's 659 MPs are from ethnic minority groups, he added. Commission for racial equality chief Trevor Phillips argued on Sunday the time had come for such shortlists. That came after it emerged that one of Britain's most ethnically diverse constituency, West Ham, was to get a women-only shortlist for the next election following an NEC ruling.\n \n Mr Phillips said changes to the race relations legislation might allow political parties to reserve seats for under-represented groups. For example in West Ham, this might allow only women and minorities to seek to be candidates. "If we get to the other side of the general election and find that minorities are not represented, we have to say that after 20 or 30 years of talking about this, we cannot go on the same way," he said. He added: "It would be terribly disappointing if in the least white constituency [West Ham] in the whole of Europe we didn't have a minority

candidate." Appearing on BBC Radio 4's Today programme, Mr Malik, who is himself running for the seat of Dewsbury, acknowledged that so far women-only shortlists had failed to deliver a boost in the number of ethnic minority candidates.\n \n But he argued: "I do think that there currently things that parties can do and which they aren't doing... for example setting targets to ensure that existing democratic structures are more reflective." Labour MP Diane Abbot, who backs Mr Phillips\' proposal of shortlists, said she had been elected along with three other ethnic minority MPs - Keith Vaz, Paul Boateng and Bernie Grant - in 1987 but it took another 10 years before another black woman was able to win a seat. That was a rate of progress Ms Abbott described as "painful". "I am a little older than Shahid and served on the National Executive Committee in the 1990s - I was the first black person on the NEC ... crossing our fingers and hoping we are going to get more black and Asian MP hasn't worked," she said. "The shortlist strategy works for women and I believe that it can be made to work for black and Asian people." On Tuesday Labour chairman Ian McCartney said his party was "ambitious" to improve black and Asian representation. "We haven't ruled out all-black shortlists and welcome a debate in the party about this," he said.\n'

'UK \needs true immigration data'\n \n A former Home Office minister has called for an independent body to be set up to monitor UK immigration.\n \n Barbara Roche said an organisation should monitor and publish figures and be independent of government. She said this would counter "so-called independent" groups like Migration Watch, which she described as an anti-immigration body posing as independent. Migration Watch says it is not against all immigration and the government already publishes accurate figures. Sir Andrew Green, chairman of the organisation, says there is no need for an independent body because Office of National Statistics data are accurate. He says he opposes large-scale immigration "both on the grounds of overcrowding and culture".\n \n He said: "For example, over the next 20 years one household in three will be due to immigration. "We are already more overcrowded than India and we are four times more overcrowded than France." Ms Roche, Labour MP for Hornsey and Wood Green, believes legal migration is something we should welcome. She said her proposals mean "we wouldn't have so-called independent experts, like Migration Watch, who come into this debate from an anti-immigration point of view." She went on: "What I would like to see is there being a body which actually looked at the figures, published them, and was independent of government. "I think that would go a long way to allaying some of the fears that are sometimes whipped up during this debate."\n'

'Call to save manufacturing jobs\n \n The Trades Union Congress (TUC) is calling on the government to stem job losses in manufacturing firms by reviewing the help it gives companies.\n \n The TUC said in its submission before the Budget that action is needed because of 105,000 jobs lost from the sector over the last year. It calls for better pensions, child care provision and decent wages. The 36-page

submission also urges the government to examine support other European countries provide to industry. TUC General Secretary Brendan Barber called for "a commitment to policies that will make a real difference to the lives of working people."
"Greater investment in childcare strategies and the people delivering that childcare will increases the options available to working parents," he said. "A commitment to our public services and manufacturing sector ensures that we can continue to compete on a global level and deliver the frontline services that this country needs." He also called for "practical measures" to help pensioners, especially women who he said "are most likely to retire in poverty". The submission also calls for decent wages and training for people working in the manufacturing sector.

'MP's shock at long lost daughter
Labour MP Stephen Pound has told of his shock and his joy on learning he fathered a daughter when he was "out of control" during the "wild" 1960s.
Lucy, 36, tracked down Mr Pound after her birth mother told her he was known as "Precious" at school. Despite being told he was a "nutter who died at sea" she found a reference to Mr Pound on Friends Reunited. Lucy's mother kept her pregnancy secret and had her adopted at four weeks. Mr Pound found out about it last year. That happened when Lucy wrote to his Parliamentary office saying she believed he was her father.
Mr Pound, who is now married with two other teenage children, said: "I was a nutter and I was a sailor but I wasn't dead." He said his first reaction was to wonder if he was victim of a "set-up" but he then realised all the dates fitted. "It was a pretty horrible thing with a pretty happy ending. I felt guilt in the marrow of my bones," said the Ealing North MP. "I don't blame Lucy's mother. I was pretty much out of control. I was 18 when she was conceived. It was a different time - it was pretty wild."

Mr Pound says he rang Lucy as soon as he got her letter, they met up and have been in contact ever since. Describing that first meeting at London's King Cross station, he said: "The earth went from under my feet ... We were walking across the Euston Road and I took her arm to take her across and there was an immediate connection. "We were finishing each other sentences." Lucy, who was adopted by a family from Essex, wants to remain anonymous for the sake of her adoptive father and her children. Mr Pound said his wife Maggie had been fully supportive and their two children Emily, 16, and Pelham, 14 were pleased to have an older sister.

"What the election should really be about?
A general election is the best chance most pressure groups get to make a real impact on government policy. Here is how six leading lobbies plan to make sure their cause is being debated ahead of an expected Spring poll.
We've called for the state pension to be increased from Â£79.60 to the pensioner credit guarantee level of Â£105.45. That's what we're calling for. Many pensioners are disadvantaged by the current system. If we've got one in five pensioners below the poverty line, we've got to make it more generous or have these people living in poverty.
We've drawn up a pensioners' manifesto. This will be sent to each of the candidates in the 659 constituencies. They will be asked which of

the top five issues, including the pension issue, they would support. Once we've got their responses we will publish the results within the constituencies and nationally as well. It's our way of putting the politicians on notice. We are trying to get across the fact that there are 11m voters over 60 in the country, they are more likely to vote than other sections of society and thirdly they are true swing voters. Before 1997 most pensioners voted Conservative. In 1997 and 2001 they voted Labour. But there is no guarantee they will vote for a Labour government this time around. They cannot take that vote for granted.

Pensions generally will certainly be a big election issue even though the government has postponed the publication of Adair Turner's full report into the issue. He said the UK had one of the least generous pensions systems in the developed world.

That the government takes seriously the impact of aviation on the environment.

We haven't worked out specific plans but I imagine we will lobby political parties and incumbent MPs. Various local groups will do that in their particular areas and we will provide a national briefing. We don't have any large demonstrations planned but they can't be ruled out.

It is hard to say whether we will be successful. We have got the issue in the public consciousness to an extent, but it is difficult to say whether an election will raise its importance in the public mind or whether it will be pushed out by big issues like Iraq.

Repealing the Hunting Bill.

We are challenging the use of the Parliament Act 1949 in a High Court action. We are hoping to hear in the New Year. Whichever way the court rules the other side will appeal so we expect it to fall plumb in electioneering time. When the ban comes into force on 18 February we will be going to the European Court because no compensation is being paid. So there's a lot of legal territory to go. We are trying to engage with the ministers by demonstrating and talking. Whatever intelligence we get we will try to turn up and speak to whoever it is. (Rural affairs minister) Alun Michael has avoided us and cancelled engagements so that makes it difficult. It is not intimidatory - on the whole it is groups of angry housewives. Of course there is an element of shouting because people are angry but there is no violence because that does not achieve anything.

It will fall plumb in the run up to the most important general election Tony Blair will ever face. It's exactly what the prime minister did not want. He wanted the issue off the table until after the election.

People using live animals as targets for sport both here and abroad. The reason for including abroad is because of trophy hunting. It is another sort of form of shooting for sport. The principle is the same whether it's a tiger or a pheasant.

We will widely publicise what's happening in relation to trophy hunting. We will publicise the darker aspects of the target animal industry the UK. We will seek to get pledges from individual MPs and would-be MPs saying that they are against the use of animals as targets for sports. We would like the support of political parties but I think a general election is very much to do with pledges MPs make to their electors. With hunting we had many MPs who were happy to say they were against

it.\n \n I think what we will get is a very real climbing up the agenda. Whether or not we will get a ban I am not sure. But it will mobilise public opinion. Everything we do will reduce animal suffering and in time that will lead to a ban.\n \n The issue that we think is the most important for this election is choice. The language of consumerism is very commonplace in government and across the political spectrum. Choice as an ideology is beginning to be the privatisation of this decade. It's become an issue in itself but what's really missing from the debate is the consumer's choice in that. Choice is not choice at all if all you have to choose from is two failing schools. We have seen so many pensions mis-selling scandals and in the pensions industry there's a maximum of choice but a minimum quality in that. We want choice on the consumer's terms - that means clear and accessible information to operate that choice.\n \n Firstly, we have our website. It features our campaigns and changes every day. Secondly through our 700,000 members who communicate with us. Thirdly through the media and also what we will be doing is holding a pre-election conference. We will invite the opinion formers, MPs, journalists and others. The idea is that we open up communications between members of the public and the politicians.\n \n I think we will be successful. It's very much the language being used by the main political parties. Politicians on all sides are very sensitive to this issue they want to be seen to be responding to this issue.\n \n Fuel duty is a large part of operational costs for road haulage workers. We have been hearing about this proposed increase of 1.92p per litre that Gordon Brown has been postponing and postponing. Tuppence does not sound like a great deal but every year if you operate one vehicle that's an increase of about Â£750. If you're running 10 vehicles it's obviously 10 times that.\n \n If fuel duty does rise we will be absolutely horrified. There will be a huge effect throughout the industry and I would not be surprised if you see widespread demonstrations. What it will mean is there will be a number of firms going out of business.\n \n We will continue to do what we have always done we keep the issue in the trade press. Regrettably it's one of those stories that it is getting harder to get into the national press. Whatever we do, the public don't like lorries - they see us as a complaining minority. But they don't realise that when you see a car on the road it is probably going to work, when you see a lorry it's already at work.\n"]

CLUSTER 2

['Label withdraws McFadden\\'s video\n \n The new video of former Westlife singer Brian McFadden has been pulled after a Dublin school complained about being associated with his song Irish Son.\n \n St Fintian\\'s High School says it is clearly identified in the video, while McFadden never went there. McFadden makes claims that he was beaten at his own school in the song\\'s lyrics, saying it had "cell blocks". The performer\\'s record label Sony BMG has withdrawn the

video and issued replacements to television stations. The label believed the school name was fictitious until they received the complaint. They have said the reference to the school was unintentional and coincidental.\n \n The head of Christian Brothers' school St Fintian\''s, Richard Fogarty, said the video implied that the 24-year-old pop star had attended his school and was abused there. "The school has always treated its pupils with respect," Mr Fogarty said in a statement. McFadden makes specific references to the Christian Brothers in his song, but did not attend St Fintian\''s. Corporal punishment was outlawed in Irish schools in 1982 when McFadden was two years old.\n \n McFadden, whose debut solo album is released next week, has said that every song is autobiographical and "a true story". Alcoholism and domestic violence are among the other topics dealt with in his songs, half of which have been written with Robbie Williams\'' former collaborator, Guy Chambers. McFadden, who quit chart-topping group Westlife in March, went to number one in September with his first solo single Real To Me. He enjoyed 12 chart-toppers with the boy band before parting company with them.\n'

'Bangkok film festival battles on\n \n Organisers of the third Bangkok International Film Festival have been determined to carry on with this year\''s event despite the ravages of the Asian tsunami disaster.\n \n The festivities have been scaled down, red carpets have been mothballed and profits from ticket sales are being donated to the tsunami relief fund. Apart from this, however, the festival has continued as originally planned. "When the disaster happened, we naturally asked ourselves if we should cancel," said the festival\''s executive director, Craig Prater. "The decision was made that we would continue, but that the focus would change.\n \n "Our premieres became fundraisers, the opening night was visibly toned down and 10% of every ticket sold goes to the disaster fund," he continued. "But we feel like we\''ve turned a page. We\''ve acknowledged our position, and now it\''s business as usual."\n \n Mr Prater\''s sentiments were echoed by Christine Rush, director of the festival\''s sister event, the Bangkok Film Market. "We have been careful to be respectful of the Thai nation\''s recovery from this terrible disaster," she said.\n \n "However, the nation\''s recovery very much depends upon it continuing its economic life. "We are encouraging attendees to keep the victims in mind and aggressively support the aid organisations helping out," she added. Given the conspicuously low audience figures, that support may be more symbolic than significant. However, poor ticket sales have less to do with any fallout from the disaster and more to do with administrative and promotional difficulties.\n \n The eclectic programme contains a wide range of titles, including such Oscar hopefuls as Ray Charles biopic Ray and Mike Leigh\''s abortion drama Vera Drake. That said, the absence of Thai subtitles on most film prints might have proved offputting to local cinemagoers.\n \n On a more positive note, the celebrity guests attending the event have generated huge media interest in this busy capital city. Problems securing air transport, the clash with last Sunday\''s Golden Globes

ceremony in Los Angeles and other factors have forced some big names to cancel.\n \n But the festival has still attracted such Hollywood directors as Oliver Stone and Joel Schumacher, alongside such screen stars as Michael Douglas, Jeremy Irons and Jean-Claude Van Damme. Irons, whose film *Being Julia* is in competition for the festival's Golden Kinnaree awards, said he had wondered whether he should attend after December's catastrophe. "I questioned whether I should come when such a tragedy had happened," the British actor said. "I thought it might be in rather bad taste. But then I spoke to the organisers and I was encouraged to let life go on. "Now I've come here I'm very glad. What impresses me about Thai culture is they get over these things far quicker than we do in the West," he continued. "Basically I was invited here by Thailand. If they still want you to come, you should."\n \n Now in its third year, Bangkok is a relatively recent addition to the hectic film festival calendar. And while it has yet to attain the stature of more established events held in Hong Kong, Tokyo and Pusan in South Korea, Mr Prater predicts it will soon become their equal. "It's grown awfully fast in three years - maybe too fast," he concedes. "We've had some bumps in the road, but that's only because the festival is still a baby. "I have no doubt in my mind that in five years, this will be the top Asian film festival in the world." The Golden Kinnarees will be awarded on Friday. The festival itself runs until 24 January.\n'

'Fox '\too reliant on reality TV'\n \n The head of US TV network Fox has admitted the broadcaster had relied too heavily on reality TV shows such as the poor-rating *Who's Your Daddy*. \n \n Chief executive Gail Berman said "in the case of this fall we drifted to too much on the unscripted side". The series *Who's Your Daddy*, where a young woman tries to pick her natural father for a cash prize caused outrage from adoption groups and rated badly. Last season, Fox's prime-time audience fell by 600,000 to 5.9 million. Ms Berman said: "I think the audience expects loud things from Fox. Sometimes they work, and sometimes they don't." \n \n *Who's Your Daddy*, the first episode of which was shown on 3 January, pulled in a disappointing audience of 6.3 million, according to the Nielsen ratings system. Five other episodes of the show had also been filmed will be dropped from Fox's schedules, Ms Berman said. She was predicting a drop in ratings even for some of the network's established reality shows, such as *American Idol*, which is due to start its fourth series this week. Fox had unveiled a new strategy last year promising to launch new shows every season, including the traditionally quiet summer season. Though that had met with a poor reception, Ms Berman said "there's no question that the audience, in our mind, is ready, willing and able to accept new programming in the summer". Fox has changed this plan, launching new shows in May instead of June. One of the new shows will be the animated series *American Dad*, made by Seth MacFarlane, the creator of *Family Guy*. That series, after becoming a hit on DVD, is also set to return with new episodes.\n'

'India to deport Bollywood actress\n \n India has ordered the

deportation of Iranian-born model and actress Negar Khan to Norway after saying she was working illegally on her visa.\n \n Khan has had raunchy roles in music videos and Bollywood films over the past two years. The distressed actress told media she was being driven straight to the airport after a routine appointment at the Bombay immigration office. Immigration officials said she had been warned about her visa last year.\n \n Khan told an Indian news channel on her mobile phone: "They are not even letting me go home to get my bags. I have no clothes or money on me. "They did not give me any notice. I don't know why they are taking me away like this. There is nothing wrong with my visa. It says I work here... They are not allowing me to even call a lawyer." Bipin Bihari, deputy police chief in Bombay, said Khan was being sent back to Norway, from which she holds a passport, on the first available flight. "She was on a visitor's visa but she has engaged in several commercial ventures," he said. "We issued notices last year, in view of which she had gone back to Norway, but she has returned again this year to work." One of Khan's biggest successes was the steamy Hindi song Chadti Jawani (Rising Youth). Khan was involved in controversy when topless photos alleged to be of her appeared in a Norwegian magazine.\n'

'Film star Fox behind theatre bid\n \n Leading actor Edward Fox is one of the men behind plans to reopen a Swansea theatre thought to be Wales'\ oldest.\n \n The star of film and TV hits Day of the Jackal and Edward and Mrs Simpson has joined forces with Swansea-born actor-director Terry Palmer. They will set up a charity to raise money to buy the Palace Theatre, and hope to reopen it to audiences in summer 2006. It is estimated that Â£500,000 is needed to buy the Grade II listed building. The Palace Theatre, which will be known as The Pavilion Repertory Theatre, was sold to an undisclosed bidder at auction 13 months ago for Â£340,000. Before that it had been used as a nightclub for 10 years.\n \n Oscar-winner Sir Anthony Hopkins, who was born in nearby Port Talbot, made his professional debut there, and Charlie Chaplin and Lily Langtry have also trodden its boards. Fox told the South Wales Evening Post that after visiting the theatre he was surprised by its beauty. "A lot of money needs to be spent on it, but all the bones are there - it just needs a beautiful skin," he said.\n \n "To call it a jewel is not an over-estimation. There are not many theatres left like this. "The hope for the immediate future is that we can attract enough people to come forward to allow it to prosper," he told the newspaper.\n \n Fox first found fame in the 1960s. One of his most memorable roles was as an assassin in the Day of the Jackal in 1973, and he won a TV Bafta as the Prince of Wales in 1980 for Edward and Mrs Simpson. He was brought on board to save the theatre by Mr Palmer, a long-standing friend of 40 years. "Two years ago I decided to return to Swansea and do something for the city," said Mr Palmer. He said he was excited by the idea of running the Palace as a charity where young actors could develop, and with an annual Shakespeare festival in the summer. "If in five years' time the theatre can function in all areas without me, I shall be well satisfied having helped to create a vital living theatre for the

city," he added. Keith Poulton, a business adviser from Swansea, has already committed an undisclosed amount of money to the project. "We need at least £75,000 to secure a deposit on the building and we've only got two months to do that," said Mr Poulton. "A few benefactors have expressed their interest and Mr Fox is going to give it an awful lot of time," he added. Last year, roads around the theatre were cordoned off when masonry from its roof fell onto a car, but engineers said they were happy that it was safe.\n']

CLUSTER 3

['Wall Street cool to eBay's profit\n \n Shares in online auction house eBay fell 9.8% in after-hours trade on Wednesday, after its quarterly profits failed to meet market expectations.\n \n Despite seeing net profits rise by 44% to \$205.4m (£110m) during October to December, from \$142m a year earlier, Wall Street had expected more. eBay stock fell to \$92.9 in after-hours trade, from a \$103.05 end on Nasdaq. eBay's net revenue for the quarter rose to \$935.8m from \$648.4m, boosted by growth at its PayPal payment service.\n \n Excluding special items, eBay's profit was 33 cents a share, but analysts had expected 34 cents.\n \n "I think Wall Street has gotten a bit ahead of eBay this quarter and for the 2005 year." said Janco Partners analyst Martin Pyykkonen. For 2004 as a whole, eBay earned \$778.2m on sales of \$3.27bn. eBay president and chief executive Meg Whitman called 2004 an "outstanding success" that generated "tremendous momentum" for 2005. "I'm more confident than ever that the decisions and investments we're making today will ensure a bright future for the company and our community of users around the world," she said. eBay now forecasts 2005 revenue of \$4.2bn to \$4.35bn and earnings excluding items of \$1.48 to \$1.52 per share. Analysts had previously estimated that eBay would achieve 2005 revenues of \$4.37bn and earnings of \$1.62 per share, excluding items.\n'

'Russian oil merger excludes Yukos\n \n The merger of Russian gas giant Gazprom and oil firm Rosneft is to go ahead, but will not include Yugansk, which was controversially bought last year.\n \n The merger, backed by Russian authorities, will allow foreigners to trade in Gazprom shares. Gazprom chief Alexei Miller confirmed Rosneft-owned Yugansk was not part of the deal and will instead be spun off. Under the agreement, the state will get a controlling share of Gazprom in exchange for Rosneft. The state wanted to control Gazprom before allowing foreigners to trade. Speaking on NTV television, which is controlled by Gazprom, Mr Miller added that Yugansk, which was swallowed up by Rosneft late last year, will operate as a separate, state-owned oil firm headed by current Rosneft chief Sergei Bogdanchikov. According to reports from Russian News Agency Interfax, the deal should go through in the next two to three months.\n \n "Obtaining majority control over Gazprom is the beginning of the liberalisation of the market in Gazprom shares," Mr Miller added. By

opening up trading in Gazprom to foreigners, the firm will become a top emerging market play for traders. Currently, foreigners can only trade in Gazprom via a small issue of London-listed proxy shares. "This is positive news for the international investment community," Global Asset Management investment chief David Smith said. "The majority of investors are going to be happy," he added.\n \n However, analysts were disappointed that Yugansk would not be included in the deal. "Yugansk is a heavy cashflow generator and would have been a much better asset for Gazprom," Renaissance Capital energy analyst Adam Landes told Reuters news agency. But he said the latest development was simply an interim step to allow foreigners to trade in Gazprom. "Ultimately and industrially, Gazprom needs Yugansk," he added. Analysts said the deal would give Gazprom control of 8% of Russia\'s total oil production, an improvement on its current 2.5%, but still far less than the 20% share it would have gained had it also taken over Yugansk. However, the merged group will still remain outside Russia\'s top five oil producers - led by Lukoil with 11% of the market , followed by TNK-BP which is half owned by BP, and Surgutneftegaz. Instead, the merged Gazprom-Rosneft group will rank alongside Sibneft with 7% of the market. Yugansk was sold to a little-known shell company in a disputed auction in December, following what many thought was a politically-motivated attack on Yukos. The shell company was then snapped up by Rosneft. Yukos unsuccessfully sought to halt the auction by applying for bankruptcy through the US courts. The unit was auctioned by Russian authorities to help pay off a \$27.5bn back-tax bill.\n'

'Fannie Mae \'should restate books\'\n US mortgage company Fannie Mae should restate its earnings, a move that is likely to put a billion-dollar dent in its accounts, watchdogs have said.\n The Securities & Exchange Commission accused Fannie Mae of using techniques that "did not comply in material respects" with accounting standards. Fannie Mae last month warned that some records were incorrect. The other main US mortgage firm Freddie Mac restated earnings by \$5bn (£2.6bn) last year after a probe of its books. The SEC\'s comments are likely to increase pressure on Congress to strengthen supervision of Fannie Mae and Freddie Mac.\n The two firms are key parts of the US financial system and effectively underwrite the mortgage market, financing nearly half of all American house purchases and dealing actively in bonds and other financial instruments. The investigation of Freddie Mac in June 2003 sparked concerns about the wider health of the industry and raised questionsmarks over the role of the Office of Federal Housing Enterprise Oversight (OFHEO), the industry\'s main regulator. Having been pricked into action, the OFHEO turned its attention to Fannie May and in September this year said that the firm had tweaked its books to spread earnings more smoothly across quarters and play down the amount of risk it had taken on. The SEC found similar problems. The watchdog\'s chief accountant Donald Nicolaisen said that "Fannie Mae\'s methodology of assessing, measuring and documenting hedge

ineffectiveness was inadequate and was not supported" by generally accepted accounting principles.\n'

'India\'s Deccan seals \$1.8bn deal\n Air Deccan has ordered 30 Airbus A320 planes in a \$1.8bn (£931m) deal as India\'s first low-cost airline expands in the fast-growing domestic market.\n Air Deccan was set up last year and wants to lure travellers away from the railway network and pricier rivals. The potential of the Indian market has attracted attention at home and abroad. Beer magnate Vijay Mallya recently set up Kingfisher Airlines, while UK entrepreneur Richard Branson has said he is keen to start a local operation.\n The country has a population of more than a billion people and many observers feel that it is underserved by airlines. Recently however, the booming economy has boosted personal spending power and helped swell the middle classes and the corporate sector. India\'s government has given its backing to cheaper and more accessible air travel. "The days of flying being a symbol of only maharajas or the rich are over," the minister for civil aviation Praful Patel said earlier.

Infrastructure is being built to handle the expected increase in demand and on Tuesday, Agence France Presse reported that a group led by Germany\'s Siemens won the contract to build a private airport near Bangalore. India\'s airports authority and the state government will own 13% each of the finished transport hub. For its part, Air Deccan, set up by army officer and silk farmer Gorur Gopinath, plans to increase its fleet to 60 aircraft within five years. To help finance the expansion the company may sell a 25% stake to an investor for about \$50m. When it was set up the firm offered tickets that were 50% cheaper than other Indian airlines. It said it was basing its business model on European firms such as Ireland\'s Ryanair.\n'

'Worldcom ex-boss launches defence\n Lawyers defending former WorldCom chief Bernie Ebbers against a battery of fraud charges have called a company whistleblower as their first witness.\n Cynthia Cooper, WorldCom\'s ex-head of internal accounting, alerted directors to irregular accounting practices at the US telecoms giant in 2002. Her warnings led to the collapse of the firm following the discovery of an \$11bn (£5.7bn) accounting fraud. Mr Ebbers has pleaded not guilty to charges of fraud and conspiracy.\n Prosecution lawyers have argued that Mr Ebbers orchestrated a series of accounting tricks at WorldCom, ordering employees to hide expenses and inflate revenues to meet Wall Street earnings estimates. But Ms Cooper, who now runs her own consulting business, told a jury in New York on Wednesday that external auditors Arthur Andersen had approved WorldCom\'s accounting in early 2001 and 2002. She said Andersen had given a "green light" to the procedures and practices used by WorldCom. Mr Ebber\'s lawyers have said he was unaware of the fraud, arguing that auditors did not alert him to any problems.\n Ms Cooper also said that during shareholder meetings Mr Ebbers often passed over technical questions to the company\'s finance chief, giving only "brief" answers himself. The prosecution\'s star witness, former WorldCom financial chief Scott Sullivan, has said that Mr Ebbers ordered accounting adjustments at

the firm, telling him to "hit our books". However, Ms Cooper said Mr Sullivan had not mentioned "anything uncomfortable" about WorldCom's accounting during a 2001 audit committee meeting. Mr Ebbers could face a jail sentence of 85 years if convicted of all the charges he is facing. WorldCom emerged from bankruptcy protection in 2004, and is now known as MCI. Last week, MCI agreed to a buyout by Verizon Communications in a deal valued at \$6.75bn.\n']

CLUSTER 4

['Movie body hits peer-to-peer nets\n \n The movie industry has struck out at file-sharing networks with another round of lawsuits in the US.\n \n The Motion Picture Association of America (MPAA) also said it had succeeded in getting a network called LokiTorrent closed down. It is the latest network which uses the peer-to-peer system called BitTorrent to be hit by the MPAA. The MPAA began its legal campaign against operators of similar networks across four continents in December. A Dallas court agreed that Hollywood lawyers would be allowed access to LokiTorrent's server records which could let them single out those who were sharing files illegally. In October 2004, the site had provided links to more than 30,000 files. The action came after the operators of LokiTorrent agreed a settlement with the MPAA. A stark message has appeared on the site from the MPAA warning "You can click, but you can't hide". In BitTorrent systems, server sites do not host the files being shared. They host links, called "trackers" that direct people to others that have it instead.\n \n As well as filing an unspecified number of file suits across the US, the MPAA said it had given operators that host eDonkey servers "take down" notices. Hollywood studios are aggressively clamping down on file-sharers who it says infringe copyright laws by copying films and TV programmes then share the files online. But it is now targeting the operators of BitTorrent networks themselves. It has filed 100 lawsuits against operators of BitTorrent server sites since December. The strategy of hitting those who run the servers which link to copyrighted material is intended to stunt file-sharers' ability to swap content using BitTorrent systems. The film industry says the black market for illegally copied videos and DVDs already costs them billions every year and it is worried that illegal file-sharing is adding to their losses. In December, the legal action claimed its most high-profile victim. The popular Suprnova.org website was forced to close, and others like Phoenix Torrent followed soon after.\n'

'Be careful how you code\n \n A new European directive could put software writers at risk of legal action, warns former programmer and technology analyst Bill Thompson.\n \n If it gets its way, the Dutch government will conclude its presidency of the European Union by pushing through a controversial measure that has been rejected by the European Parliament, lacks majority support from national governments and will leave millions of European citizens in legal limbo and facing

the possibility of court cases against them. If the new law was about border controls, defence or even the new constitution, then our TV screens would be full of experts agonising over the impact on our daily lives. Sadly for those who will be directly affected, the controversy concerns the patenting of computer programs, a topic that may excite the bloggers, campaigning groups and technical press but does not obsess Middle Britain. After all, how much fuss can you generate about the Directive on the Patentability of Computer-Implemented Inventions, and the way it amends Article 52 of the 1973 European Patent Convention? Yet if the new directive is nodded through at the next meeting of one of the EU's ministerial councils, as seems likely, it will allow programs to be patented in Europe just as they are in the US. Many observers of the computing scene, including myself, think the results will be disastrous for small companies, innovative programmers and the free and open source software movement. It will let large companies patent all sorts of ideas and give legal force to those who want to limit their competitors' use of really obvious ideas.

In the US you cannot build a system that stores customer credit card details so that they can pay without having to re-enter them unless Amazon lets you, because they hold the patent on "one-click" online purchase. It is a small invention, but Amazon made it to the patent office first and now owns it.

We are relatively free from this sort of thing over here, but perhaps not for long. The new proposals go back to 2002, although argument about patentability of software and computer-implemented inventions has been going on since at least the mid-1980s. They have come to a head now after a year in which proposals were made, endorsed by the Council of Ministers, radically modified by the European Parliament and then re-presented in their original form. Some national governments seem to be aware of the problems. Poland has rejected the proposal and Germany's main political parties have opposed it, but there is not enough opposition to guarantee their rejection. Early in December the British government held a consultation meeting with those who had commented on the proposals. Science Minister Lord Sainsbury went along to listen and outline the UK position, but according to those present, it was embarrassing to see how little the minister and his officials actually understood the issues concerned.

The draft Directive is being put through the council as what is called an "A" item and can only be approved or rejected. No discussion or amendment is allowed. So why should we be worried? First, there is the abuse of the democratic process involved in disregarding the views of the parliament and abandoning all of their carefully argued amendments. This goes to the heart of the European project, and even those who do not care about software or patents should be worried. If coders are treated like this today, who is to say that it will not be you tomorrow?

More directly, once software patents are granted then any programmer will have to worry that the code they are writing is infringing someone else's patent. This is not about stealing software, as code is already protected by copyright. Patents are not copyright, but

something much stronger. A patent gives the owner the right to stop anyone else using their invention, even if the other person invented it separately. I have never, to my shame, managed to read Lord Byron's Childe Harold's Pilgrimage. If it was pointed out that one of my articles contained a substantial chunk of the poem then I could defend myself in court by claiming that I had simply made it up and it was coincidence. The same does not hold for a patent. If I sit down this afternoon and write a brilliant graphics compression routine and it happens to be the same as the LZW algorithm used in GIF files, then I am in trouble under patent law, at least in the US. Coincidence is no defence. The proposed directive is supported by many of the major software companies, but this is hardly surprising since most of them are US-based and they have already had to cope with a legal environment that allows patents. They have legal departments and, more crucially, patents of their own which they can trade or cross-license with other patent holders.\n\nEven this system breaks down, of course, as Microsoft found out last year when they initially lost a case brought by Eolas which claimed that Internet Explorer (and other browsers) infringed an Eolas patent. That one was eventually thrown out, but only after months of uncertainty and millions of dollars. But small companies, and the free and open software movement do not have any patents to trade. Much of the really useful software we use every day, programs like the Apache web server, the GNU/Linux operating system and the fearsomely popular Firefox browser, is developed outside company structures by people who do not have legal departments to check for patent infringements. The damage to software will not happen overnight, of course. If the directive goes through it has to be written into national laws and then there will be a steady stream of legal actions against small companies and open source products. Eventually someone will decide to attack Linux directly, probably with some secret funding from one or two large players. The new directive will limit innovation by forcing programmers to spend time checking for patent infringements or simply avoiding working in potentially competitive areas. And it will damage Europe's computer industry. We can only hope that the Council of Ministers has the integrity and strength to reject this bad law.\n\nBill Thompson is a regular commentator on the BBC World Service programme Go Digital.\n'

'What's next for next-gen consoles?\n\nThe next generation of video games consoles are in development but what will the new machines mean for games firms and consumers? We may not know when they will be released, what they will be called or even what they will be able to do but one thing is certain - they are coming. Sony, Microsoft and Nintendo are all expected to release new machines in the next 18 months. The details of PlayStation 3, Xbox 2 (codename Xenon) and Nintendo's so-called Revolution are still to be finalised but developers are having to work on titles for the new machines regardless.\n\n"We know maybe what the PS3 will do, but we can only guess," said Rory Armes, studio general manager for video game giant Electronic Arts in Europe. "It's a horrendous effort in the first

year," he admitted. Microsoft had delivered development kits to EA, said Mr Armes, but he said the company was still waiting on Sony and Nintendo to send kits. Although the details may not be nailed down, Mr Armes said EA was beginning to get a sense of the capabilities of the new machines. "The rumours are that PlayStation 3 will have a little more under the hood [than Xbox 2]," he said.\n \n "Microsoft is obviously a software company first and foremost, while Sony has more experience in hardware. I think Sony will be able to push more into a box at cost."\n \n What is certain is that the new machines will provide great leaps in processing and graphical power. It is also likely that they will contain convergence technologies to make the machines more of an entertainment hub. At the Consumer Electronics Show in Las Vegas Microsoft boss Bill Gates said Xbox 2 would be part of his vision of the digital lifestyle. While short on detail, he painted a picture of a machine that would complement a PC and offer "great video gaming but video gaming for a broader set of people, more communications, more media, more connectivity". Sony is thought to have a similar vision for PS3 while Nintendo remain focused on a machine purely for games.\n \n Until it is clear precisely what the new machines can do, developers are working on a first round of titles to harness the new horsepower. Gary Dunn, development director at Codemasters, said the company had a central technology group which was at the forefront of preparing for the next generation of consoles. "We are working on new libraries of effects. A lot of the major techniques are already out there and in use in PC gaming, such as pixel shaders and normal mapping."\n \n Mr Dunn said he expected the introduction of real-world physics to be a major part of the new consoles. "We want to increase that level of immersion and realism in gaming to people can lose themselves in a game." In the first year at least, developers said gamers should not expect games which harness the full potential of the machines. Graphical spit and polish and better physics in line with the capabilities of current high-end graphics cards for PCs should be expected. Simon Gardner, president of Climax\ 's Action Studio, said: "It\ 's definitely an exciting time. We want to give more freedom to the player. We want to give players an emotional connection to the characters they play. "The environments will be much more believable and dramatic, growing and changing as you play. "There will be a breadth of effects, more involving worlds to play in.\n \n "It\ 's a bit like being an artist and being given a bigger canvas and a smaller brush. We\ 're being given more tools. "For the average consumer, we can get things of a more filmic quality." Gerhard Florin, head of EA in Europe, said gamers should expect titles that blur the line between films and video games. Many will be sceptical - gamers were given similar predictions during the last transition of console hardware - but this time it would seem to be more likely. "PS3 will provide graphics indistinguishable from movies," said Mr Florin. He said the distribution method for games would also change radically in the next round of consoles. "A gamer could buy a starter disc for 10 euros. When he goes home he goes online and he could buy AI and levels

as you go.\n "It's much smarter if you can get levels as you go." Mr Armes warned that developers still had to learn how to tell stories effectively in the medium. "In some ways we are trying to forget about the hardware, go in the opposite direction. We have been very bad at letting technology design our creativity. "What we have to do as a company is start ignoring the technology and learning our craft in telling stories." Mr Gardner agreed: "We can throw more polygons around and have better AI but if it doesn't make for a better game then that's not very useful." Developers will certainly have the tools with the new machines, but how they employ them is still to be decided.\n"

"Electronics firms eye plasma deal\n Consumer electronics giants Hitachi and Matsushita Electric are joining forces to share and develop technology for flat screen televisions.\n The tie-up comes as the world's top producers are having to contend with falling prices and intense competition. The two Japanese companies will collaborate in research & development, production, marketing and licensing. They said the agreement would enable the two companies to expand the plasma display TV market globally.\n Plasma display panels are used for large, thin TVs which are replacing old-style televisions. The display market for high-definition televisions is split between models using plasma display panels and others - manufactured by the likes of Sony and Samsung - using liquid-crystal displays (LCDs). The deal will enable Hitachi and Matsushita, which makes Panasonic brand products, to develop new technology and improve their competitiveness. Hitachi recently announced a deal to buy plasma display technology from rival Fujitsu in an effort to strengthen its presence in the market.\n Separately, Fujitsu announced on Monday that it is quitting the LCD panel market by transferring its operations in the area to Japanese manufacturer Sharp. Sharp will inherit staff, manufacturing facilities and intellectual property from Fujitsu. The plasma panel market has seen rapid consolidation in recent months as the price of consumer electronic goods and components has fallen. Samsung Electronics and Sony are among other companies working together to reduce costs and speed up new product development.\n"

'Peer-to-peer nets \'here to stay'\n Peer-to-peer (P2P) networks are here to stay, and are on the verge of being exploited by commercial media firms, says a panel of industry experts.\n Once several high-profile legal cases against file-sharers are resolved this year, firms will be very keen to try and make money from P2P technology. The expert panel probed the future of P2P at the Consumer Electronics Show in Las Vegas earlier in January. The first convictions for P2P piracy were handed out in the US in January. William Trowbridge and Michael Chicoine pleaded guilty to charges that they infringed copyright by illegally sharing music, movies and software. Since the first successful file-sharing network Napster was forced to close down, the entertainment industry has been nervous and critical of P2P technology, blaming it for falling sales and piracy. But that is going to change very soon, according to the panel.\n

The music and film industries have started some big legal cases against owners of legitimate P2P networks - which are not illegal in themselves - and of individuals accused of distributing pirated content over networks. But they have slowly realised that P2P is a good way to distribute content, said Travis Kalanick, founder and chairman of P2P network Red Swoosh, and soon they are all going to want a slice of it. They are just waiting to come up with "business models" that work for them, which includes digital rights management and copy-protection standards.\n \n But, until the legal actions are resolved, experimentation with P2P cannot not happen, said Michael Weiss, president of StreamCast Networks. Remembering the furore around VCRs when they first came out, Mr Weiss said: "Old media always tries to stop new media. "When they can't stop it, they try to control it. Then they figure out how to make money and they always make a lot of money." Once the courts decided that the VCR in itself was not an illegal technology, the film studios turned it into an extremely lucrative business. In August 2004, the San Francisco-based US Court of Appeals ruled in favour of Grokster and StreamCast, two file-sharing networks. The court said they were essentially in the same position that Sony was in the 1980s VCR battle, and said that the networks themselves could not be deemed as illegal.\n \n P2P networks usually do not rely on dedicated servers for the transfer of files. Instead it uses direct connections between computers - or clients. There are now many different types of P2P systems than work in different ways. P2P nets can be used to share any kind of file, like photos, free software, licensed music and any other digital content. The BBC has already decided to embrace the technology. It aims to offer most of its own programmes for download this year and it will use P2P technology to distribute them. The files would be locked seven days after a programme aired making rights management easier to control. But the technology is still demonised and misunderstood by many. The global entertainment industry says more than 2.6 billion copyrighted music files are downloaded every month, and about half a million films are downloaded a day. Legal music download services, like Apple iTunes, Napster, have rushed into the music marketplace to try and lure file-sharers away from free content. Sales of legally-downloaded songs grew tenfold in 2004, with 200 million tracks bought online in the US and Europe in 12 months, the IFPI reported this week. But such download services are very different from P2P networks, not least because of the financial aspect.\n \n There are several money-spinning models that could turn P2P into a golden egg for commercial entertainment companies. Paid-for-pass-along, in which firms receive money each time a file is shared, along with various DRM solutions and advertiser-based options are all being considered. "We see there are going to be different models for commoditising P2P," said Marc Morgenstern, vice president of anti-piracy firm Overpeer.\n \n "Consumers are hungry for it and we will discover new models together," agreed Mr Morgenstern. But many net users will continue to ignore the entertainment industry's potential controlling grip on

content and P2P technology by continuing to use it for their own creations. Unsigned bands, for example, use P2P networks to distribute their music effectively, which also draws the attention of record companies looking for new artists to sign. "Increasingly, what you are seeing on P2P is consumer-created content," said Derek Broes, from Microsoft. "They will probably play an increasing role in helping P2P spread," he said. Looking into P2P's future, file sharing is just the beginning for P2P networks, as far as Mr Broes is concerned. "Once some of these issues are resolved, you are going to see aggressive movement to protect content, but also in ways that are unimaginable now," he said. "File-sharing is the tip of the iceberg."\\n']

```
category_dict = {
    0: 'Sports',
    1: 'Politics',
    2: 'Entertainment',
    3: 'Business',
    4: 'Technology'
}
# mapping cluster labels to categories
clustered_data['Category'] =
clustered_data['Category'].map(category_dict)

clustered_data.head()

{"summary": {"\n    \"name\": \"clustered_data\", \n    \"rows\": 2127,\n    \"fields\": [\n        {\n            \"column\": \"Text\", \n            \"properties\": {\n                \"dtype\": \"string\", \n                \"num_unique_values\": 2127,\n                \"samples\": [\n                    \"Blair congratulates Bush on win\\n \\n Tony Blair has said he looks\n                    forward to continuing his strong relationship with George Bush and\n                    working with him during his second term as president.\\n \\n Mr Bush's\n                    re-election came at a crucial time for a world that was \\"fractured,\n                    divided and uncertain\\\", Mr Blair said. It had to be brought\n                    together, he added, saying action was needed on poverty, the Middle\n                    East and the conditions on which terrorists prey. Mr Blair said states\n                    had to work with the US to fight global terrorism.\\n \\n But there\n                    was a need to recognise it would not be defeated by \\"military might\n                    alone but also by demonstrating the strength of our common values\\\"\n                    he added. Solving the conflict in the Middle East was the world's\n                    single most \\"pressing political challenge\\\" of the present day,\n                    Mr Blair warned. The prime minister also urged Europe and the US to\n                    \\"build anew their alliance\\.\\n \\n \\"All of us in positions\n                    of leadership, not just President Bush, have a responsibility to rise\n                    to this challenge. It is urgent that we do so.\\\" Mr Blair also paid\n                    tribute to Democrat John Kerry's campaign, saying he had helped make\n                    the presidential election \\"a true celebration of American\n                    \"]\n            }\n        }\n    ]\n}
```

democracy\\\". The election of the US president was significant for the world but particularly so for Britain because of its special relationship, he added. Earlier Tory leader Michael Howard sent Mr Bush his \\\"warmest congratulations\\\", saying: \\\"We look to the president to be a unifying force for those all over the world who share our determination to defend freedom.\\\" Liberal Democrat leader Charles Kennedy welcomed the fact there had been a quick conclusion to the election, unlike in 2000.\\n \\n Mr Bush's first task was to \\\"rebuild a sense of domestic purpose\\\" within the US, he said. Mr Kennedy said: \\\"Internationally, it is to be hoped that a second term will see a more sensitive approach to relations with long-standing allies, not least for the global efforts to combat terrorism.\\\" Lib Dem foreign affairs spokesman Menzies Campbell said a win by Mr Kerry would have given Mr Blair the chance of a fresh start, adding it was almost as if there was an \\\"umbilical cord\\\" between Mr Bush and the UK premier.\\n \\n \\\"Europeans must hope that his administration will be much more multilateral in character, and that he will act swiftly to rebuild the Atlantic partnership which is so vital to security. \\\"Iraq will remain an issue of potential division for some time to come.\\\" Even before the result became clear, Mr Blair was being urged to push for action on climate change at his first meeting with whichever candidate won.\\n \\n Liberal Democrat environment spokesman Norman Baker underlined the issue of global warming during a Commons debate on Anglo-American relations on Wednesday. Environment Secretary Margaret Beckett has said the US will act on global warming despite George Bush's refusal to sign up to the Kyoto protocol on carbon emissions. Public opinion would force change, she told BBC news. But Myron Ebell, an adviser on climate change to President Bush, has said there would be no change in the US stance and rejected the threat of climate change. He claimed the US was the only country with independent scientists.\\n\\n \\\"China Aviation seeks rescue deal\\n \\n Scandal-hit jet fuel supplier China Aviation Oil has offered to repay its creditors \$220m (\\u00c2\\u00a3117m) of the \$550m it lost on trading in oil futures.\\n \\n The firm said it hoped to pay \$100m now and another \$120m over eight years. With assets of \$200m and liabilities totalling \$648m, it needs creditors' backing for the offer to avoid going into bankruptcy. The trading scandal is the biggest to hit Singapore since the \$1.2bn collapse of Barings Bank in 1995. Chen Jiulin, chief executive of China Aviation Oil (CAO), was arrested by at Changi Airport by Singapore police on 8 December. He was returning from China, where he had headed when CAO announced its trading debacle in late-November. The firm had been betting heavily on a fall in the price of oil during October, but prices rose sharply instead.\\n \\n Among the creditors whose backing CAO needs for its restructuring plan are banking giants such as Barclay's Capital and Sumitomo Mitsui, as well as South Korean firm SK Energy. Of the immediate payment, the firm - China's biggest jet fuel supplier - said it would be paying \$30m out of its own resources. The rest would come from its parent company, China Aviation Oil Holding Company in


```

<IPython.core.display.HTML object>

Saving news_article_labels.csv to news_article_labels.csv

labels.shape
(2127, 1)

# checking the unique labels
labels['Label'].unique()

array(['Politics', 'Sports', 'Technology', 'Entertainment',
       'Business'],
      dtype=object)

clustered_data['Actual Category'] = labels['Label'].values

print(classification_report(clustered_data['Actual Category'],
                           clustered_data['Category']))

          precision    recall   f1-score   support
Business        0.96     0.93     0.95      503
Entertainment    0.96     0.95     0.95      369
Politics         0.96     0.95     0.95      403
Sports           0.98     0.99     0.99      505
Technology       0.91     0.97     0.93      347

accuracy           0.96      --      0.96      2127
macro avg         0.95     0.96     0.96      2127
weighted avg      0.96     0.96     0.96      2127

incorrect_category_data = clustered_data[clustered_data['Actual
Category'] != clustered_data['Category']].copy()
incorrect_category_data.shape
(91, 3)

incorrect_category_data.head()

{"summary": "{\n  \"name\": \"incorrect_category_data\", \n  \"rows\": 91,\n  \"fields\": [\n    {\n      \"column\": \"Text\", \n      \"properties\": {\n        \"dtype\": \"string\", \n        \"num_unique_values\": 91,\n        \"samples\": [\n          \"Wal-Mart is sued over rude lyrics\\n \\n The parents of a 13-year-old girl\n          are suing US supermarket giant Wal-Mart over a CD by rock group\n          Evanescence that contains swear words.\\n \\n The lawsuit, filed in\n          Washington County, alleges Wal-Mart deceived customers by not putting\n          warning labels on the cover. Trevin Skeens alleges Wal-Mart knew of\n          the offending word because it had censored it on its music sales\n          website. Wal-Mart said it was investigating the claims but had no\n          warning labels.\"]\n      }\n    }\n  ]\n}"}}

```

plans to pull the CD. Wal-Mart has a policy of not stocking CDs which carry parental advisory labels. Mr Skeens said he bought the Anywhere But Home CD for his daughter and was shocked to hear the swearing when it was played in their car.\n \n \"I don't want any other families to get this, expecting it to be clean. It needs to be removed from the shelves to prevent other children from hearing it,\" said Mr Skeens of Brownsville. The lawsuit seeks to force Wal-Mart to censor the music or remove it from its stores in Maryland. It also seeks damages of up to \$74,500 (\u00c2\u00a338,660) for every customer who bought the CD at Maryland Wal-Marts, and also naming record label Wind-Up Records and distributor BMG Entertainment in the legal action.

\"While Wal-Mart sets high standards, it would not be possible to eliminate every image, word or topic that an individual might find objectionable,\" Wal-Mart spokesman Guy Whitcomb told the Herald-Mail of Hagerstown.\n ,\n

Campbell to be Lions consultant\n Former government communications chief Alastair Campbell will act as a media consultant to Sir Clive Woodward's 2005 Lions on their tour to New Zealand.\n Campbell, who left Downing Street earlier this year, will advise on media strategy before and during the tour. \"I hope I can contribute to the planning and preparation, and to ensuring the media and public get the most out of the tour itself,\" he said. \"I am also looking forward to going out for the later stages of the tour.\" Woodward's decision to call in Prime Minister Tony Blair's former spin doctor springs from the deterioration in media relations on the last Lions tour of Australia in 2001, when New Zealander Graham Henry was the head coach. The furore surrounding the newspaper diaries of Matt Dawson and Austin Healey was compounded by other disillusioned players venting their frustration through the media. \"The Lions is a massive media event,\" said Woodward, who will be the head coach. \"There will be a huge level of interest from the travelling media, the fans that will go out in their thousands and the New Zealand public. \"We need to have the strategy and processes in place to deal with the pressures that will bring. \"[Alastair] will act as an advisor both in the build up to and on the tour itself. His role is to work closely with not only myself but (tour manager) Bill Beaumont, (media manager) Louisa Cheetham and (team manager) Louise Ramsay.\" Campbell is due to resume working for the government in the new year in the build-up to an anticipated May general election. The Lions leave for New Zealand on 24 May, with the first Test match against the All Blacks in Christchurch on 25 June.\n

UK TV channel rapped for CSI ad\n TV channel Five has been criticised for sending \"offensive\" and \"threatening\" advertising material to viewers for a new show about murder scene scientists.\n Five mailed thousands of fake dossiers including photos of murder victims and an e-mail suggesting the recipient was being stalked by a serial killer. Following complaints, the Advertising Standards Authority contacted Five to cease promotion of crime show CSI:NY. Five admitted it had sent out 55,000 promotion packs but had now stopped. The

promotion material was sent in brown envelope of the type used by investigators in the series, a spin-off from the highly successful CSI: Crime Scene Investigation series, which also runs on Five. The pack also features pictures of forensic evidence from a crime scene and a wanted poster, which did have a CSI:NY logo printed in large at the bottom.\n \n Five said it had received 100 complaints but that it had been surprised at the reaction because it was \\\"obvious this material is promoting a drama\\\". A Five spokesman said: \\\"In light of the efforts we have made to make the nature of the contents so transparent we are surprised a very small minority of recipients have mistaken it as anything else. \\\"Everyone who was sent this promotion has expressed an interest in receiving details about this particular genre of programming on various websites. \\\"We have also received emails and calls from recipients praising the originality and imagination of the campaign.\\\"\\n\\n],\n

```

  "semantic_type": "\\",\n      "description": "\\\"\\n      }\n    },\n      {\n        "column": "\\\"Category\\\",\\n        "properties": {\n          "dtype": "\\\"category\\\",\\n          "num_unique_values": 5,\n          "samples": [\n            "\\\"Technology\\\",\\n\n            "\\\"Politics\\\",\\n            "\\\"Business\\\"\\n          ],\n          "semantic_type": "\\\",\\n            "description": "\\\"\\n      }\n    },\n      {\n        "column": "\\\"Actual Category\\\",\\n        "properties": {\n          "dtype": "\\\"category\\\",\\n          "num_unique_values": 5,\n          "samples": [\n            "\\\"Sports\\\",\\n            "\\\"Business\\\",\\n            "\\\"Technology\\\"\\n          ],\n          "semantic_type": "\\\",\\n            "description": "\\\"\\n      }\n    }\n  ]\n},\n  "type": "dataframe", "variable_name": "incorrect_category_data"
}

idx = 24

print('Distance from Actual Category')
print(cdist(embedding_matrix[idx].reshape(1,-1),
kmeans.cluster_centers_=[[2]], "euclidean")[0,0])

print('Distance from Predicted Category')
print(cdist(embedding_matrix[idx].reshape(1,-1),
kmeans.cluster_centers_=[[3]], "euclidean")[0,0])

Distance from Actual Category
0.9545187652210279
Distance from Predicted Category
1.0357046825280756

idx = 45

print('Distance from Actual Category')
print(cdist(embedding_matrix[idx].reshape(1,-1),
kmeans.cluster_centers_=[[2]], "euclidean")[0,0])

```

```
print('Distance from Predicted Category')
print(cdist(embedding_matrix[idx].reshape(1,-1),
kmeans.cluster_centers_[[4]], "euclidean")[0,0])

Distance from Actual Category
1.019115605584521
Distance from Predicted Category
1.0080196953598326
```