A1. Consider using a computer to decrypt a general mixed-alphabet cryptogram by brute force. Suppose the computer can check 10^9 decryption alphabets per second (including the time required to decide whether the result is sensible or not).

(a) How long will it take for the computer to check all 26! possible alphabets? Give your answer in the most appropriate time units.

Solution: Using the calculator, we find that $26! \approx 4.0329 \times 10^{26}$. We divide this by 10^9 to get 4.0329×10^{17} seconds.

We divide this by 3600 to get 1.12025×10^{14} hours.

We divide this by 24 to get 4.6677×10^{12} days.

We divide this by 365.25 to get 12 779 535 234.8 years. (The precision was carried to the limit of the calculator through all calculations.) This is about 12.8 billion years.

For comparison, the Earth is thought to have formed about 4.5 billion years ago.

(b) Suppose that we are able to guess (correctly) the ciphertext equivalents of the letters `t`, `h`, and `e`. Once these letters are fixed, how many decryption alphabets will we have to check? How long will it take?

Solution: With three letters fixed, we have $23! \approx 2.5852 \times 10^{22}$ ways to arrange the other 23.

We divide this by 10^9 to get 2.5852×10^{13} seconds.

We divide this by 3600 to get 7 181 115 760.8 hours.

We divide this by 24 to get 299 213 156.7 days.

We divide this by 365.25 to get about 819 201 years.

(c) Now suppose we know the ciphertext equivalents of `t`, `h`, and `e`, and we are quite confident that the four least common letters in the cryptogram are the cipertext equivalents of `q`, `j`, `x`, and `z` (but we don't know which of these four is which). How many decryption alphabets will we have to check? How long will it take?

Solution: Three letters are fixed, and the number of ways to arrange the other 23 is equal to 4! (the number of wasy to arrange the equivalents of `q`, `j`, `x`, and `z`) times 19! (the number of way to arrange the rest). There are $4! \times 19! \approx 2.9195 \times 10^{19}$ possible alphabets.

At 10^9 tries per second, we could explore them all in

2 919 482 410 seconds, or
810 967 hours, or
33 790 days, or
92.5 years.

(d) Finally, suppose that we do frequency counts on a great deal of text and find the following:

- There is one letter that is clearly the ciphertext equivalent of `e`;
- There are seven other high-frequency letters, which must correspond (in some order) to `t`, `a`, `i`, `n`, `o`, `s`, and `r`;
- There are four medium-frequency letters, which must correspond (in some order) to `h`, `l`, `d`, and `c`;
- There are five very low-frequency letters, which must correspond (in some order) to `k`, `j`, `q`, `z`, and `x`.

How many decryption alphabets will we have to check? How long will it take?

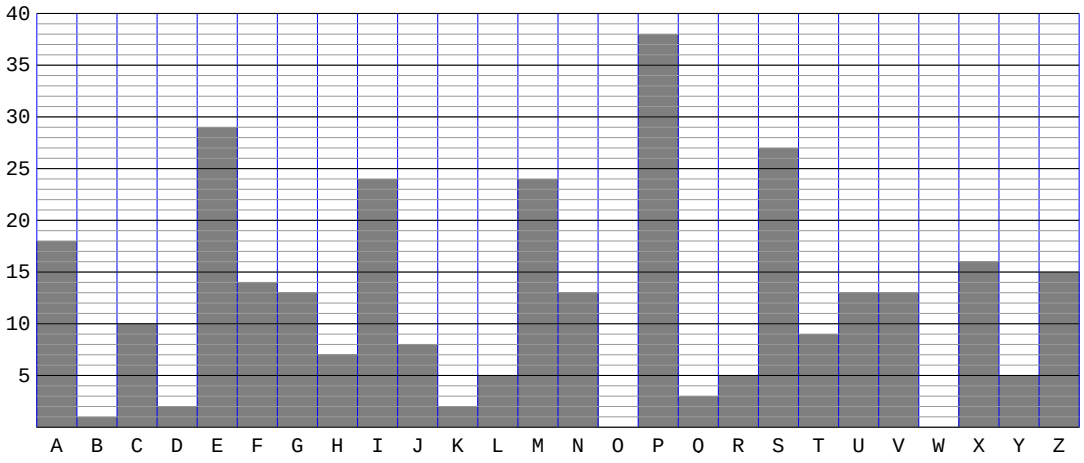
Solution: This time, only one letter is fixed. The high-frequency letters can be arranged in 7! ways, the medium-frequency letters in 4! ways, the low-frequency letters in 5! ways, and the remaining letters in 9! ways.

There are $7! \times 4! \times 5! \times 9! \approx 5.2673 \times 10^{12}$ possible alphabets.

At 10^9 tries per second, we could explore them all in 5 267 seconds, which is just under 88 minutes.

B1. Pick one of the cryptograms on the handout Cryptograms #6 and cryptanalyze it. Provide a short explanation of how you reconstructed the encryption or decryption alphabet. You don't need to decrypt the entire message (unless you want to); the first hundred letters will do.

I chose the Cryptogram 3. First I did a frequency histogram. Here's the picture



It seems very likely that plaintext `e` is ciphertext `P`, and that plaintext `t` is one of `E`, `I`, `M`, or `S`. A pattern search turned up three occurrences of the string `ETP`, so it seems likely that plaintext `t` is ciphertext `E` and plaintext `h` is ciphertext `T`.

I did a pattern search for the word `groceries` (using the repeated `r` and the two `e`'s). Two matching strings turned up: `UESBPETPJ` and `UZIVPZMPX`. The letter frequencies in the latter string seemed more reasonable, so I started constructing an alphabet using `groceries` $\mapsto$ `UZIVPZMPX`.

With these letters in place, I was able to recognize some words: from `co__e__ie__ces` I got `n` and `v`, from `e__cee__ing`, I got `x` and `d`, from `h__ving`, I got `a`, and so on.

Here's the completed alphabet:

Ciphertext:	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Plaintext:	n	k	d	x	t	u	y	f	o	m	q	b	i	l		e	v	p	a	h	g	c		s	w	r

Here's the decrypted text:

> Personally, while I am exceedingly fond of America and grateful for its many conveniences, I am not quite so slavishly accepting. Take the matter of having your groceries bagged for you. I appreciate the gesture and all, but when you come down to it, what does it actually get you except the leisure to stand and watch your groceries being bagged? It's not as if it buys you some quality time.

B2. Pick one of the cryptograms on the handout Cryptograms #7 and cryptanalyze it. Provide a short explanation of how you determined the keyword. Give a decryption of at least the first hundred letters of the cryptogram.

I chose Cryptogram 2. I used Microsoft Word to split the letters into five cosets. They are

First letters:

```
PIRXSVRZVVRLCRVXYVRMURCZEVRVWIUZZERLVJZWIJVXT
RGKKYVFPSCFRFJDYPKJKVKKIDTCRZFIZVUAFCJTSUFDDJ
YIJACZZEIFYFREVXUYZFITIZEVZPFJWYVKRKUXWGIZZR
```

Second letters:

```
MLDIMAGVIVIZQDZMMIBIIAQAIWVKXQMABOSKVQAWQPIIM
TQPPBECUZGZQKBQMOIBPVPCMISIJKUQDIIIVIQMTCXIGP
GMUCWMVCIXINVGVWGWANQCMPWZVETQIQPWKPLMQCBKWVV
```

Third letters:

```
YGGXIZYYSJTXYKOYTJYJRYTYZTQKGCXJXZKQJTUXKKZIV
GKKUYZMOOHKJGUTUXXYKZIXTTCTRGUGKNIIVITYKMGQKA
VGCYXJZSTGJOOJJTCROGIRGGOKGGYUTTGMQKRYTXOBTEM
```

Fourth letters:

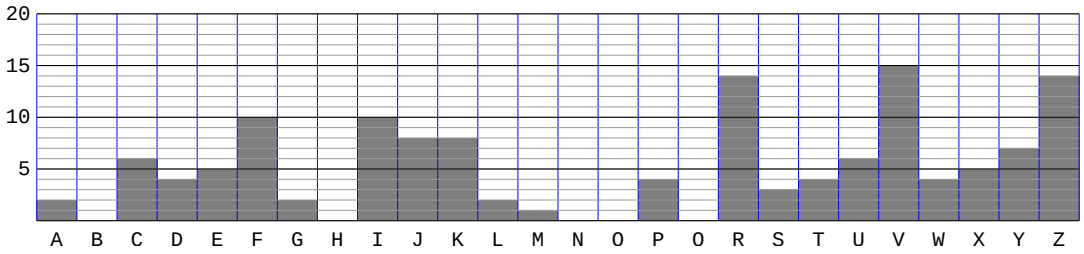
```
AFUKVVTALHKLTHAVPAHVPLNPBIWZWVLHFVHIVNDNZJPVA
YJZBMOOULLPDYTKULAVAPLFVYLKVZUUZLORVRHZHSYLFY
LSHAPWIIKYAUYLPHHZVTHAUCUZUYCUFNCVATLAKLZPVA
```

Fifth letters:

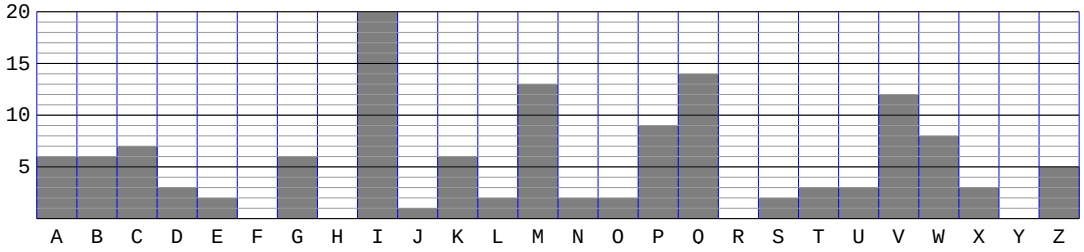
```
XLMXFWTKOWLTAATPKAEKPEAZKEBHXGWEBFURKAGXBKGGN
MXXZEKFWYYLAXRMETBYPXGTKHEIILWZFTXLEBVBGRMLXM
KBLZYTRXIMHXHIGGKOGXGNWXMMWABHMPXUHBTHTTMLYA
```

Here are the frequency histograms from the five cosets.

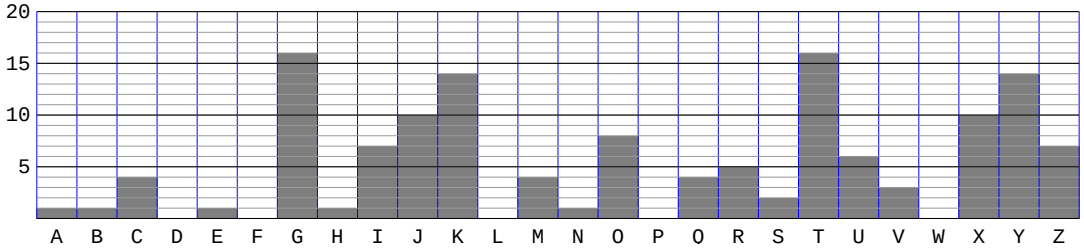
First letters:



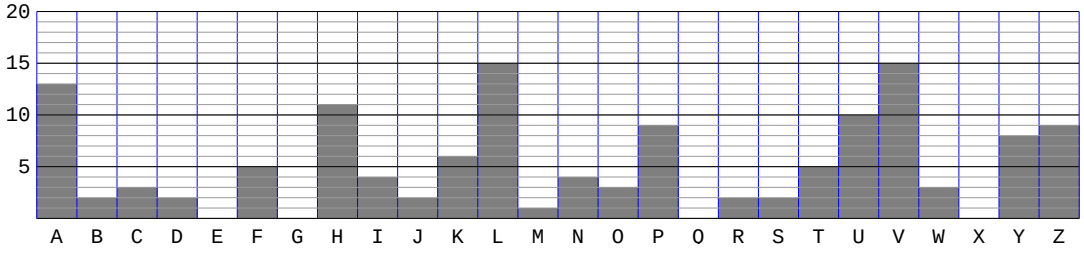
Second letters:



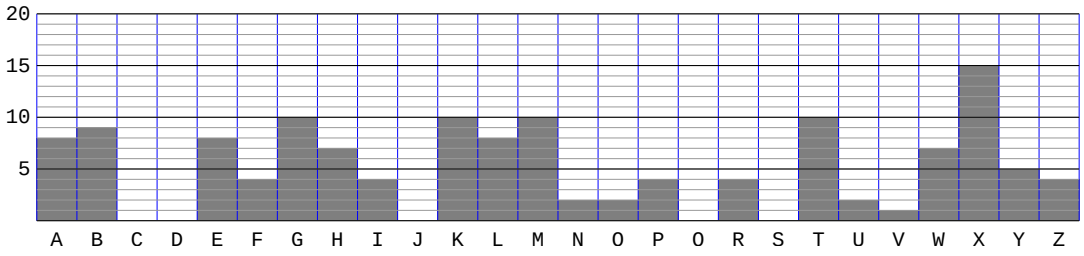
Third letters:



Fourth letters:



Fifth letters:



From the shapes of the histograms, we can make good guesses at the keys for the five individual shift ciphers. The first is `a` $\mapsto$ `R`, the second is `a` $\mapsto$ `I`, the third is `a` $\mapsto$ `G`, the fourth is `a` $\mapsto$ `H`. The fifth is a little ambiguous, but *tta* $\mapsto$ `T` would work, and it is also the only letter that completes a sensible keyword. We guess that the keyword is `RIGHT`.

Here is a decryption of the text:

> Yesterday's avant-garde becomes today's mainstream. Even dada and surrealism have a heritage. So when I read that Salvador Dali was selling his signature on blank pieces of paper, I wondered, "Is Dali trying to make a buck by endorsing his own forgeries? Is he creating a conceptual art piece?" These thoughts flew through my mind briefly before I said, "Who cares?" To my mind the only great artists of the twentieth century are Norman Rockwell and Pablo Picasso. Mondrian gives me a headache. Jackson Pollack is inaccessible and ugly. Opart makes my eyes hurt. Hyperrealism was just glorified paint-by-number, and popart had too fine an irony depending on Andy Warhol's vision of American culture. And I have no interest in Andy Warhol's vision of anything. We have to go back to the Middle Ages to find a pure artistic vision of anything.