(5) Theorem. Every finite p-group can be embedded in a group of unitriangular matrices over $\mathbb{F}_p$.

Proof. The theorem will be proved in the following steps.

(1) $|GL_n(p)| = \prod_{i=0}^{n-1}(p^n - p^i)$

By definition $GL_n(p)$ is the set of all non-singular $n \times n$ matrices over the field $\mathbb{F}_p$. It is sufficient to have a set of n linearly independent row vectors, each of length n, to construct such a matrix. Without any restriction there are p^n choices for a row vector, with p choices for each element. Here, since we are constructing a non-singular matrix there are $p^n - 1$ choices for the first row after excluding the $\bar{0}$ vector. After the first one is chosen there are p multiples (including $\bar{0}$) of this row which must not be chosen for the second row to maintain linear independence. Hence the second row has $p^n - p$ choices.

For the third row we will have to discard the p multiples of both the first and the second row to ensure linear independence, and hence can be chosen in $(p^n - p^2)$ ways. Continuing likewise the last row must not be a multiple of the the first $n-1$ rows and hence has $(p^n - p^{n-1})$ options. Now (1) can be established using product rule for counting.

(2) $UT_n(p)$ is a Sylow-p subgroup of $GL_n(p)$.

$UT_n(p)$ is the group of upper triangular $n \times n$ matrices over $\mathbb{F}_p$ with 1s in the diagonal. Since such matrices have determinant 1 (the product of diagonals) we have $UT_n(p) \subset GL_n(p)$, and hence $UT_n(p) \leq GL_n(p)$. Now observe that fixing the diagonals and the lower diagonal entries leaves us $1+2+\cdots+(n-1) = n(n-1)/2$ spaces to be filled by arbitrary elements. Since each element has p choices we have,

$$|UT_n(p)| = p^{\frac{n(n-1)}{2}}$$

But also observe that,

$$|GL_n(p)| = \prod_{i=0}^{n-1}(p^n - p^i) = \prod_{i=0}^{n-1} p^i \cdot (p^{n-i} - 1) = p^{\frac{n(n-1)}{2}} \cdot \prod_{i=0}^{n-1}(p^{n-i} - 1)$$

But $p \nmid \prod_{i=0}^{n-1}(p^{n-i}-1)$, hence a Sylow-$p$ subgroup of $GL_n(p)$ have order $p^{\frac{n(n-1)}{2}}$.

(3) S_n can be embedded in $GL_n(p)$.

Consider V, a vector space over $\mathbb{F}_p$ of dimension n. An automorphism of V is an invertible map $f : V \to V$. We know that all such maps uniquely determines a non-singular matrix T with entries from $\mathbb{F}_p$. Conversely, every such non-singular T defines an invertible map $f : V \to V$. Moreover, composition of such maps represents multiplication of the corresponding matrices and the identity map corresponds to the identity matrix. Thus the immediate isomorphism $Aut(V) \cong GL_n(p)$ can be established.

Now fix $\{v_1, v_2, \cdots v_n\}$ as a basis of V. To define an automorphism of V it is sufficient to specify the images of the basis vectors. Consider $\theta : S_n \to Aut(V)$, $\sigma \mapsto f$, where $f(v_i) = v_{\sigma(i)}$; f is an automorphism as the image of the basis set under f is itself and hence is linearly independent. Since every σ induced permutation of the basis set define a distinct automorphism we have θ as a monomorphism, or θ is one-one. Therefore θ is embedded in $Aut(V)$ and hence in $GL_n(p)$.

(4) Every finite p-group is embedded in a group of unitriangular matrices over $\mathbb{F}_p$.

Let $|G| = p^k$. By Cayley's theorem G is embedded in a subgroup of S_{p^k}. By (3), S_{p^k} is embedded in $GL_{p^k}(p)$, hence, G is isomorphic to a p-subgroup of $GL_{p^k}(p)$. But we have as a direct consequence of Sylow theorems that every p-subgroup of a group must be contained in one of its Sylow-p subgroups. Therefore G is embedded in some Sylow-p subgroup Q of $GL_n(p)$. But from (2) and Sylow's theorems we have Q and $UT_n(p)$ as conjugates, and hence isomorphic. So in conclusion, G is embedded in $UT_n(p)$.